2009

Identification of Tree Species in Mixed Conifer Forests Using Laser Altimetry

Agus Suratno

IDENTIFICATION OF TREE SPECIES IN MIXED CONIFER

FORESTS USING LASER ALTIMETRY

By

AGUS SURATNO

B.Sc. The University of Montana, Missoula, MT, 2002
M.Sc. The University of Montana, Missoula, MT, 2005

Dissertation

presented in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy
in Forestry

The University of Montana
Missoula, MT

August, 2009

Approved by:

Perry Brown, Associate Provost for Graduate Education
Graduate School

Dr. Carl Seielstad, Chair
School of Forestry and Conservation

Dr. Lloyd Queen
School of Forestry and Conservation

Dr. Woodam Chung
School of Forestry and Conservation

Dr. David Affleck
School of Forestry and Conservation

Dr. Anna Klene
Department of Geography

Suratno, Agus. Ph.D August, 2009 Forestry

Identification of Tree Species in Mixed Conifer Forests Using Laser Altimetry

Chair: Dr. Carl Seielstad

Airborne laser scanning data is used to discriminate between Douglas-fir (*Pseudotsuga menziesii*), ponderosa pine (*Pinus ponderosa*), lodgepole pine (*Pinus contorta*), and western larch (*Larix occidentalis*) in a mixed coniferous forest of western Montana, USA. Laser-derived structural and intensity variables are investigated to classify tree species at individual tree and plot-level dominant species. Linear Discriminant Analysis is applied to discriminate between species, and in combination with Maximum Likelihood Classification, used to map species across the landscape. One-way ANOVA tests indicate that proportions of first and single returns and mean intensities are significantly different between species (*p-value* < 0.001) at both individual tree and plot-dominant species levels. A single variable in the Linear Discriminant Analysis (LDA), e.g., mean or standard deviation intensity, can produce classification accuracy ranging from 49-61% at the dominant species level and 37-52% for individual trees. The accuracy can be improved to 95% and 68% respectively by using multiple variables, including proportions of return type, intensities, and canopy heights. Adding proportion of return-type improves classification accuracy at the dominant species level, but not for individual trees. The inclusion of both mean and standard deviation of canopy heights produce higher accuracy at both levels. Validation of the landscape classifications is performed using a stand database consisting of predominant and secondary species and with gridded, fixed-area plots. Assuming that stand homogeneity is at least within dominant species criteria (>70%), the application of intensity and canopy height variables generates a classification accuracy of 45% that is increased to 53% by including mixed species (stands without a clear dominant species) in the error analysis. A second method based on Maximum Likelihood Classification (MLC) using two layers, (1) the modified LDA-species based layer, and (2) percent canopy cover (PCC) layer, improves accuracy up to 75%. Unlike the modified LDA, in which the accuracy increases with incorporation of mixed species, the MLC produces lower accuracy (38%) when these stands are included. Both methods, the modified LDA and MLC produce best results for Douglas-fir followed by lodgepole pine and ponderosa pine, while western larch is difficult to identify. Almost without exception, the classification identifies the correct mix of species within each mixed polygon, but field data do not currently support validation of individual pixels within stands.

Acknowledgements

I am really fortunate to be part of National Center for Landscape Fire Analysis (NCLFA) in the College of Forestry and Conservation, at the University of Montana. Firstly, I would like to express my gratitude to my advisor Dr. Carl Seielstad for his invaluable contribution to the success of this work. Also, I would like to thank Professor LLoyd Queen for his support and encouragement during my study. My special thanks to Dr. Woodam Chung for his endeavor, serving as my co-advisor earlier and providing me with various support. I would like to thank the members of my supervisory committee, Dr. LLoyd Queen, Dr. Woodam Chung, Dr. Anna Klene and Dr. David Affleck for their time, advice and support during my research. Eric Rowell deserves special acknowledgement for his important work in processing and preparing laser datasets used in this study. He also coordinated the field data collection with a crew of people from NCLFA whom deserve my appreciation, including Crystal Stonesifer, Casey Teske, Erik Hakanson, Tim Wallace, Ann Hadlow, and Josh Rodriquez. I would like to thank Martin Twer, and R.J. Hannah for field assistance and Jim Riddering for his constructive comments and suggestions on portion of this work. I am grateful to Cate Crue (former NCLFA secretary) for her organizational work. I would also like to thank other NCLFA staff, including Jami Sindelar, Craig Comstock, Leana Schelvan, Lee Macholz, Limei Piao and Valentijn Hoff for their supports. I am really grateful to Saxon Holbrook and Andrew Neuschwander for their tireless efforts and time in providing IT support. I would like to thank my family in Indonesia for their endless support and prayers during my study in USA. My mother always teaches me to be patient and optimistic, and my father always encourages me to explore and experience different ways in order to be successful in every field I work. My brothers and sister always support me with their prayers and thoughtful inputs. Finally, I would like to express my greatest thanks to my wife, Akiko for her patient, faithful, confidence and encouragement during my study. Lastly, I am really proud of you Haku, my son, for being a smart and nice boy. Thank you for your inspiring smile and being talkative, motivating me to finish this work.

Tables of Contents

List of Tables

List of Figures

CHAPTER 1

Introduction to Airborne Laser Scanning Systems and Tree Species Identification

Overview

The intent of this dissertation research has been to evaluate and develop the capacity of relatively low density (< 1 return/m^2) airborne laser scanner data for discriminating between tree species in an inland Northwestern conifer forest. The work examines four common species, Douglas-fir (*Pseudotsuga menziesii*), ponderosa pine (*Pinus ponderosa*), lodgepole pine (*Pinus contorta*), and western larch (*Larix occidentalis*) and assesses the relative importance of intensity, height, and return-type metrics for classifying them in a mixed-species forest system. While most of the work focuses on species identification, exploratory analyses of both the laser system and landscape classification are central to the research. The dissertation is arranged into five chapters, starting with an exploration of laser intensity data and culminating in the application of tree-species classification at broad scale. Chapter one presents a perspective, background, and a statement of goals. Chapter two is an exploration of intensity characteristics, examines potentially confounding effects including Automatic Gain Control (AGC), and produces a simple test-case landscape classification. In this chapter, airborne laser scanning (ALS) data produced by a Leica ALS50 instrument that operates with an AGC subsystem is evaluated to identify the distinctiveness of intensity and the effects of scan angle (range) and AGC on intensity. The exploration uses samples from urban (Missoula International Airport) and natural (Lubrecht Experimental Forest/LEF) sites. Intensity normalization is also applied and its quality is evaluated. The normalized intensity data are used to classify landcover in both areas using a simple, exploratory threshold method. Chapter three presents species identification using laser-derived variables at both individual tree and plot levels. In this chapter, in addition to intensity, ALS derived structural variables are used to discriminate the four dominant species mentioned above in LEF. Several statistical analyses are performed, including Linear Discriminant Analysis (LDA) for species classification at individual tree and dominant species scales. Chapter three was published in *ISPRS Journal of Photogrammetry and Remote Sensing* (Suratno et al., 2009). Chapter four describes the application of species classification at the landscape level. Variables and the methodology used in Chapter three (LDA) are used in conjunction with a maximum likelihood classifier to catalog the entire LEF at 0.04 ha resolution. The results of the classification are evaluated using independent plot data, and against an existing stand database and a current inventory dataset. This chapter will appear in the *Proceeding of Silvilaser 2009, the 9th International Conference on Lidar Applications for Assessing Forest Ecosystems.* Chapter five identifies potential shortcomings of the methodologies for species identification and presents recommendations for future research.

Statement of problem

The innovation of airborne laser scanning (ALS) systems in forestry applications has allowed researchers and foresters to take advantages of direct three dimensional measurements of trees. The use of airborne laser scanning systems for accurate estimations of height, canopy width, and stem density are well established (Means et al., 2000; Gobakken and Naesset, 2004; Maas et al., 2008), while applications for tree species identification are relatively new. Species investigations rely on the capability of laser systems to provide both x.y,z information and intensity (Holmgren and Persson, 2004; Moffiet et al., 2005; Donoghue et al., 2007; Ørka et al., 2007). The use of intensity data in addition to structural data is relatively new. Laser intensity data have not been widely utilized for forestry applications due to uncertainties regarding signal-to-noise ratio and due to the high variability in canopy reflectivity caused by factors such as leaf morphology and condition, crown structure, and height. The latter uncertainties are ones that the passive optical remote sensing community has dealt with for more than 30 years, and that group's collective experience shows that biophysical variables of interest (in this case tree species) can often be disentangled from confounding variability (Curran, 1990; Miller et al., 1991; Kim et al., 1994; Daughtry et al., 1989). Discrimination of species of individual trees in mixed conifer forests would be a very significant advancement in both remote sensing and in forest inventory.

Currently, three primary methods have been applied to solve this problem, each originating in a different country. The Canadian method used mean laser reflectivity and variability to discriminate Jack pine from Scots pine (Schreier et al., 1985). The Scandinavian method utilized the standard deviation of the intensity of all vegetation returns and the proportion of first returns to segregate Norway spruce and Scots pine (Holmgren and Persson, 2004). The Australian method combined "vegetation permeability" with singular vegetation returns to discriminate between White Cypress pine and Poplar Box (Moffiet et al., 2005). The most recent study, conducted by Ørka et al. (2007), identified spruce, birch and aspen in Norway using the intensity of different return types. Each of these studies showed promise in classifying trees at the level of dominant species and individual tree. However, the methods have been tested only at local study sites without collectively addressing confounding effects that are known to influence intensity variations (Moffiet et al., 2005), and have not been applied to landscape classification.

One approach to species identification is to use laser data to derive species-specific crown structure measurements. In this approach, collections of laser reflections are used to define the dimensions of tree canopies which are in turn used to differentiate species with unique crown shapes. For example, Holmgren and Persson (2004) showed that spruce and pine can be discriminated in Scandinavian boreal forests by canopy shape (conical versus elliptical crowns), and similarly, in Germany, Reitberger, et al. (2006) demonstrated that spruce and beech can be identified by crown shape. An acknowledged shortcoming of the structural approach is that it requires very high data density (e.g., >3 returns/m^2), an acquisition parameter that is not often economically or logistically feasible. The limitations on data density alternatively consider other acquisitions resulting in 1-2 returns/m^2, which are practical and applicable in the large and diverse landscapes of western North America.

A second approach to species identification is to exploit differences in the reflectivity of targets vís a vís (in relation to/in regard to) analysis of laser intensity data. The latter approach is conceptually very promising, but is plagued with uncertainties regarding separation of biophysical signal from noise. The passive optical remote sensing literature clearly shows the presence of species-level differences in reflectivity at laser wavelengths (Knapp and Carter, 1998; Clark et al., 2003; Castro-Esau et al., 2006), and maps of laser intensity data depict intuitive representations of biophysical properties like landcover. Collectively, four studies have shown that dominant and individual tree species can be distinguished at a stand and tree levels using a combination of the standard deviation of laser reflection intensity, first, and singular/single returns (Schreier et al., 1985; Holmgren and Persson, 2004; Moffiet et al., 2005; Ørka et al., 2007). However, results with high accuracy are limited to classifying between conifer and deciduous tree species, and additional research is necessary to formulate effective methods for discriminating among conifer species. In addition, the intensity datasets used for these studies are produced by a system operating without automatic gain control (AGC). AGC is a method of on-the-fly intensity adjustment used by the Leica ALS50 to adjust recorded target intensity values for variations in slant range, flying height, and system AGC gain (Leica Geosystems, 2008) and it has a nonlinear effect on the output target intensity (Korpela, 2008). Complex terrain, large environmental gradients, and mixed species composition/ forest structure complicate classification of vegetation properties.

Tree species data is essential in forest ecosystem management, including inventory, forest modeling, and wildlife habitat and carbon/biomass estimation. From an economic perspective, trees of different species have different values and uses. Additionally, site productivity can be correlated to species composition and predictions of forest growth and mapping of abundance patterns often require such data. For example, habitat type classification and forest disturbance, including fire history in the northern Rockies have been distinguished using species composition at different scales (Arno, 1979; Pfister et al., 1977; Fischer and Bradley, 1987). Indeed, the importance of species information is apparent, as indicated by many well-documented studies ranging from predictions of species correlations with ecological factors such as topography and soil composition to climate change (Sollins, 1998; Hansen and Dale, 2001; Iverson and Prasad, 2001). Additionally, spatial and structural characteristics of tree species obtained using both field and remote sensing data can provide essential information for forest management and research.

Currently, essential tree and stand data, including diameter, species, canopy cover and location are obtained using field measurements combined with remote sensing data. Such data are integrated in precision forestry applications using advanced information technology (Becker, 2001). Additional estimates, including stem density and volume are derived from allometric equations and correlative functions, which are usually species specific. From this perspective, there is a critical need to develop methods of estimating forest parameters from laser measurements of height that are applicable to mixed conifer forests across a range of structures and densities. Ideally, these estimates would occur at the scale of individual tree to facilitate more precise forest management activities.

The primary research activities investigated in this dissertation are: (1) to assess the ability of low density (< 1 return/m^2) airborne laser scanning to identify tree species, and (2) to develop a methodology to map the spatial distribution of species on a large landscape (e.g., Lubrecht Experimental Forest, Montana) within the constraints of the identification methodology.

In order to meet specific goals and accomplish explicit objectives in these research areas, fundamental information of laser scanning data and species characteristics are investigated. First, laser intensity and its influencing factors are examined, because each laser system has different characteristic subsequent to quality of datasets produced (Chapter 2). Two sets of intensity data delineated from MSO and LEF samples are evaluated using general statistical metrics, including mean, standard deviation, maximum and minimum, and coefficient of variation. Intensity values are plotted against scan angle classes to observe trends. In order to assess the AGC effect, intensity normalization is performed using an existing method. Raw and normalized intensity data are compared and their variations are used to evaluate the effectiveness of normalization. A simple landscape classification is also performed on the normalized data to evaluate the potential utility of the data for classification. The next step is to examine the efficacy of laser-derived data for discriminating tree species at dominant species and individual tree levels (Chapter 3). Laser derived structural variables are used in addition to intensity metrics to discriminate between species found dominant in LEF. An accuracy assessment of results from a Linear Discriminant Analysis (LDA) is produced using field data (plots) collected in 2006 and 2007. Finally, the spatial distribution of species is generated at 0.04 ha resolution (Chapter 4). A subset of variables used in LDA at the plot level is applied to classify species using a Maximum Likelihood Classification method. Classification accuracy is calculated using two different datasets, a stand database and a permanent fixed-plot inventory dataset. Results of classification and accuracy assessment are provided for further applications and improvements.

Both laser and species characteristics are key components in this research. Therefore, the discussion is expanded to address the following questions: What are the fundamental characteristics of the laser intensity data? What are the most prominent factors contributing to its quality? Which critical information generated by laser systems is potentially useful for landscape classification? Which tree attributes are distinctive to an individual species? How can tree characteristics be identified using a laser dataset?

Although each question has a specific response, they collectively encompass the body of work described in the following four chapters. The research focuses on the classification of the four species, Douglas-fir (*Pseudotsuga menziesii*), ponderosa pine (*Pinus ponderosa*), lodgepole pine (*Pinus contorta*), and western larch (*Larix occidentalis*) located in western Montana coniferous forests. It does not attempt to evaluate differences in species spectral characteristics between data collected in the field and laboratory.

Background

Knowledge of species is important related to laser altimetry measurements of trees, because the laser's native measurement of height is often used to derive secondary forest attributes such as diameter and volume with equations that are species specific. The capability of laser altimetry or lidar (Light Detection and Ranging) in providing three dimensional measurements is advantageous compared to traditional remote sensing techniques, which only offer horizontal information. In recent years, the application of airborne laser scanning systems for obtaining landscape surface characteristics has increased significantly. Many early laser applications focused on retrieving vertical terrain distribution across landscapes (Krabill et al., 1984; Krabill et al., 1995; Kraus and Pfeifer, 1998). The high accuracy (<15 cm) of producing surface features is achievable and it could be improved depending on the system, methodology used, and topographic conditions (Krabill et al., 1984; Schreier et al., 1985; Bufton et al., 1991; Krabill et al., 1995).

Lidar applications have become increasingly important in forestry, especially for generating vegetation parameters, such as crown dimension, stem height and density.

Stand characteristics, including volume can be derived using both large (5-25 m) and small (< 2 m) footprint laser systems (Lefsky et al., 1997; Naesset, 1997, Magnussen and Boudewyn, 1998, Lefsky et al., 1999; Means, 2000). The improvement of stem delineation algorithms opens a possibility to estimate tree attributes at individual tree scales. Several tree identification methods can generate a high accuracy for separating individual tree when laser data is used in conjunction with aerial photography or near-infrared (NIR) imagery (Persson et al., 2002; Popescu et al., 2003). Persson et al (2002) used a segmentation method by creating digital canopy model (DCM) from laser data and applying a parabolic surface fitting on the NIR images to separate each stem. The method was able to detect more than 71% of individual trees in the study site and more than 91% of stem volumes were estimated when undetected small trees were removed. Popescu et al. (2003) used two variable window size (Local Maxima) techniques (square and circular windows) for delineating tree crown diameters. The R^2 values produced by these methods ranged from 0.62-0.63 for dominant trees and could be improved in deciduous stands when leaf-off visible imagery was fused with the lidar data (R^2 improves by 10%) and in the conifer trees (R^2 increases to 11%). Several researchers have investigated the possibility of using laser data alone for the individual tree detection (Popescu and Wynne, 2004; Holmgren and Persson, 2004; Rowell et al., 2009). Rowell et al.(2009) applied a combination of variable-window local maxima (LM) filtering (Popescu and Wynne, 2004) and neighborhood canopy height variance and return density (Rowell, et al. 2006) on the laser datasets used in this research. Rowell et al. (2009) found that the algorithm produced a root mean square error (RMSE) of 17.36 stems (46.8%) for overstory and intermediate trees across all structure types with much better results for certain canopy structures and crown classes.

Tree species identification is important in forest inventory subsequent to forest ecosystem management and laser datasets provide more advantages than two dimensional images. Canopy structure, intensity and combination of both have been investigated for species discrimination. Schreier et al. (1985) first opened up possibilities for tree species identification using laser intensity data. They investigated laser reflection and reflection variability measurements in mixed forests, and demonstrated the use of NIR energy from laser scanning for vegetation type identification. They found a significant difference in return intensity between pure broadleaf and coniferous forests, but the study was designed for stand level species identification, and the resolution was not high enough to identify species at a tree level due to limitations of early lidar technology.

The difference in canopy structure of individual tree species has been explored in different forest regions. In North American forests, Brandtberg et al. (2003) separated oaks, red maple and yellow poplar using leaf-off data. They suggested that variations in structural characteristics of these species were observed on laser returns. The classification accuracy of 60% was improved to 64% when a mathematical notation was introduced for grouping laser returns (Brandtberg, 2007). In Scandinavian boreal forests, Holmgren and Persson (2004) showed that spruce and pine can be discriminated in by canopy shape (conical versus elliptical crowns) while in Germany, Reitberger, et al. (2006) demonstrated that spruce and beech can be identified using the similar method (crown shape).

Species identification is possibly improved using a combination of intensity and canopy dimension measurements. Holmgren and Persson (2004) developed an approach to discriminate between pine and spruce using both lidar intensity data and shape of tree crowns on an individual tree level. They found the proportion of returns that were located above crown base height, standard deviation of the intensity of the returned pulses, and the proportion of first returns are useful variables for species identification. Their results presenting a high accuracy (95%) of classification for Scots pine and Norway spruce, again demonstrate that using laser intensity and intensity variation is promising for species identification. However, their study was conducted in the forests with sparse low vegetation, and used high density, small footprint LIDAR images that were obtained from a helicopter mounted laser system, which is not often cost-effective for forestry applications. The usability of their data analysis approach is thus uncertain when applied to datasets derived from different forest communities using different laser systems.

Similarly to Holmgren and Persson (2004), Moffiet et al. (2005) used the proportion of vegetation laser returns (vegetation permeability) and singular vegetation returns to identify species between White Cypress Pine and Poplar Box. The accuracy of their assessment was 77%, and they found the proportion of singular returns contributed most of the discriminatory power. However, this result was for the "dominant species" in the stand, and they found that a clear distinction between these two species was not always visually obvious at an individual tree level. Although their study successfully discriminated dominant tree species, they confirmed that the average return intensity and standard deviation of return intensity are highly affected not only by the reflective properties of the vegetation, but also forest structure such as tree height. They also drew a conclusion that the incident NIR pulse intensity needs to be reasonably constant over time and space in order to minimize noise effects.

The application of laser datasets for discriminating tree species remains important in forest management across different ecosystems. Two recent studies emphasize the usefulness of lidar datasets for identifying conifer and deciduous tree species (Ørka et al., 2007; Kim, 2007). Ørka et al. (2007) found that laser intensity could be useful for separating spruce, birch, and aspen trees in Norwegian forests. They also suggested that intensity variables derived from different returns affected the current overall accuracies (68-74%) up to 4% depending on the variable combination. Kim (2007) used lidar and aerial photographs to discriminate between two groups of coniferous and broadleaves species in western North American forests. Two lidar acquisitions (leaf-on and leaf-off) were used in the study and it was found that intensity metrics derived from the leaf-off data produced the best discrimination of the two species groups.

It is a well-known fact that height variation of trees affects the reflectance values of laser intensity and needs to be taken into account when used for species identification (Gaveau and Hill, 2003; St-Onge, 1999; Magnussen and Boudewyn, 1998). Tree age is another factor that could heavily affect the laser intensity because it causes different leaf structure which also controls the NIR leaf reflectance (Jacquemoud et al., 1996; Sims & Gamon, 2003; Richardson et at., 2003). While generally, the spectral properties of trees are more easily differentiated in the NIR than visible region (William, 1991), they are often affected by canopy and stem structures, and by the timing of data acquisition. For example, using hyperspectral data applied in central Sierra Nevada, Gong et al. (1997)

found that the visible bands produced stronger discriminating power than the NIR bands. They noted that the sunlit data alone produced overall accuracies greater than 91%.

Differences in laser systems can also influence intensity. Systems that record intensity directly following pulse detection by the sensor produce raw intensity, in which a weak received signal has low intensity, and a strong signal exhibits higher intensity. On the other hand, a system operating with automatic gain control (AGC), such as the datasets used for this research, generates intensity adjusted to variations in slant range, flying height and system gain value (Leica Geosystems, 2008). Several studies indicate that the adjustment has non-linear effect due to the intensity variability of targets and the AGC adjustment is often not satisfactory, especially for a target having low signal amplitude, which is dominated by noise (Miller and Wagner, 1987; Adams, 1992; Korpela, 2008). Different intensity normalizations are needed to remove or minimize such effects.

In summary, the main findings from previous studies are (1) laser-derived structural and intensity variables are useful for tree species identification, (2) distributions of intensity from vegetation do not solely represent species-specific reflectivity; (e.g., confounding effects such as tree height, tree age, moisture status and others may need to be addressed), and (3) intensity values are different between laser systems.

Project Goals

This study explores the possibility of using low density (< 1 return/m^2) laser data for species classification and investigates the following questions: Can airborne laser data be used to discriminate between four primary conifer species in mixed forests? Can species be mapped at individual tree and/or dominant species levels (at high resolution) sufficient to meet the needs of forest inventory-related applications?

With regard to the characteristics of laser data and tree species, and in order to gain a more comprehensive perspective of the classification process, the research will conduct analyses to address the following three goals:

1) Explore the fundamental characteristics of laser scanning data, including intensity and the influencing aspects (noise) and perform intensity normalization to remove/mitigate noise.
2) Develop a methodology for identifying species at individual tree and dominant species levels and provide an accuracy assessment.
3) Develop a methodology for implementing classification at landscape level and describe the species distribution within the Lubrecht Experimental Forest of The University of Montana.

Specific objectives are presented and addressed on a per-chapter basis in the subsequent chapters. Each chapter provides overview, background, methodology, results, discussion, and references for each stage of the analysis. Chapter two is primarily exploratory and chapters three and four are considered as the primary assessments that produce the species distribution map. It is important to acknowledge that alternative methods of species discrimination are not evaluated for efficacy/efficiency in this dissertation. Rather, a method is proposed, presented, and evaluated based on existing literature, flexibility, and ease-of-use. The work serves as a useful point of departure for additional investigations, as outlined in Chapter five.

References:

Adams, M.D., 2000. Lidar design, use and calibration concepts for correct environmental detection. IEEE Transactions on Robotics and Automation,16(6), 753-761.

Arno, S. F., 1979. Forest regions of Montana. USDA Forest Service, Intermountain Forest and Range Experiment Station, Research Paper, INT-218. 39pp.

Baltsavias, E.P., 1999. Airborne laser scanning: basic relations and formulas. ISPRS Journal of Photogrammetry and Remote Sensing 54 (2–3), 199–214.

Becker, G., 2001. Precision forestry in Central Europe: New perspectives for classical Management concept. In Proceedings of the first international precision forestry cooperative symposium. Seattle, 17-20 June, 2001: 7-9.

Brandtberg, T., Warner, T., Landenberger, R.E., McGraw, J.B., 2003. Detection and analysis of individual leaf-off tree crowns in small footprint, high sampling density lidar data from the eastern deciduous forest in North America. Remote Sensing of Environment 85(3), 290-303.

Brandtberg, T., 2007. Classifying individual tree species under leaf-off and leaf-on conditions using airborne lidar. ISPRS Journal of Photogrammetry and Remote Sensing, 61(5), 325-340.

Bufton, J.L., Garvin, J.B., Cavanaugh, J.F., Ramos-Izquierdo, L., Clem, T.D., Krabill, W.B., 1991. Airborne lidar for profiling of surface topography. Optical Engineering 30(1): 72-78.

Castro-Esau, K.L., Sánchez-Azofeifa, G.A., Rivard, B., Wright, S.J., Quesada, M., 2006. Variability in leaf optical properties of Mesoamerican trees and the potential for species classification. American Journal of Botany. 93:517-530.

Clark, R.N., Swayze, G. A., Wise, R., Livo, K. E., Hoefen, T.M., Kokaly, R. F., Sutley, S. J., 2003. USGS Digital Spectral Library splib05a. U.S. Geological Survey, Open File Report 03-395. http://pubs.usgs.gov/of/2003/ofr-03-395/ofr-03-395.html (accessed on January 12, 2008).

Curran, P. J., Dungan, J. L., Gholz, H.L., 1990. Exploring the relationship between reflectance, red edge chlorophyll content in slash pine. Tree Physiology. 7:33–48.

Daughtry, C. S. T., Biehl, L. L., Ranson, K. J., 1989. A new technique to measure the spectral properties of conifer needles. Remote Sensing of Environment, 27(1): 81-91.

Donoghue, D., Watt, P., Cox, N., Wilson, J., 2006. Remote sensing of species miixtures in conifer plantations using LiDAR height and intensity data. Remote Sensing of Environment 110, 509–522.

Fischer, W. C., Bradley, A.F., 1987. Fire ecology of western Montana forest habitat types. USDA Forest Service, Intermountain Forest and Range Experiment Station, Research Paper, INT-223.

Gaveau, D.L.A., Hill, R.A., 2003. Quantifying canopy height underestimation by laser pulse penetration in small-footprint airborne laser scanning data. Canadian Journal of Forest Research 29, 650-657.

Gobakken, T., Naesset, E., 2004. Estimation of diameter and basal area distributions in coniferous forest by means of airborne laser scanner data. Scandinavian Journal of Forest Research 19(6): 529-542.

Gong, P., Pu, R., Yu, B., 1997. Conifer Species Recognition: An Exploratory Analysis of In Situ Hyperspectral Data. Remote Sensing of Environment 62(2): 189-200.

Hansen, A., Dale, V., 2001. Biodiversity in US Forests under Global Climate Change. Ecosystems 4:186–199.

Holmgren, J., Persson, A., 2004. Identifying species of individual trees using airborne laser scanning. Remote Sensing of Environment 90 (4), 415-423.

Iverson, L.R., Prasad A.M., 2001. Potential changes in tree species richness and forest community types following climate change. Ecosystems 4:186–199.

Kim, M. S., Daughtry, C. S. T., Chappelle, E. W., McMurtrey, J. E., Walthall, C. L., 1994. The use of high spectral resolution bands for estimating absorbed photosynthetically active radiation. In ISPRS Sixth International Colloquium on Physical Measurements and Signatures in Remote Sensing, Val d'Ise`re, France, 17–21 January, Paris, 299–306.

Kim, S., 2007. Individual tree species identification using LIDAR-derived crown structures and intensity data. Thesis Dissertation. The University of Washington. Seattle. 137pp.

Knapp, A. K., Carter, G. A., 1998. Variability in leaf optical properties among 26 species from a broad range of habitats. American Journal of Botany, 85: 940– 946.

Korpela, I.S., 2008. Mapping of understory lichens with airborne discrete-return LiDAR data. Remote Sensing of Environment 112 (10), 3891–3897.

Krabill, W.B., Collins, J.G., Link, L.E., Swift, R.N., Butler, M.L., 1984. Airborne laser topographic mapping results. Photogrammetric Engineering & Remote Sensing 50(6): 685-694.

Krabill, W.B., Thomas, R.H., Marton, C.F., Swift, R.N., Frederick, E.B., 1984. Airborne laser topographic mapping results. Photogrammetric Engineering & Remote Sensing 50(6): 685-694.

Kraus, K. Pfeifer, N., 1998. Determination of terrain models in wooded areas with airborne laser scanner data. ISPRS Journal of Photogrammetry and Remote Sensing, 53: 193-203.

Lefsky, M.A., Cohen, W.B., Acker, S.A., Spies, T.A., Parker, G.G., Harding, D., 1997. Lidar remote sensing of forest canopy structure and related biophysical parameters at the H.J. Andrews experimental forest, Oregon, USA. In Natural Resources Management Using Remote Sensing and GIS. Greer, J.D., Editor, ASPRS, Washington, D.C.: 79-91.

Lefsky, M.A., Harding, D., Cohen, W.B., Parker, G., Shugart, H.H., 1999. Surface lidar remote sensing of basal area biomass in deciduous forests of eastern Maryland, USA. Remote Sensing of Environment, 67: 83-98.

Leica Geosystems, 2008. ALS50 Airborne Laser Scanner guide. Norcross, Georgia.

Magnussen, S., Boudewyn, P., 1998. Derivation of stand heights from airborne laser scanner data with canopy-based quantile estimators. Canadian Journal of Forest Research 28, 1016–1031.

Maas, H.G., Bienert, A., Scheller, S., Keane, E., 2008. Automatic forest inventory parameter determination from terrestrial laser scanner data. International Journal of Remote Sensing, 29(5):1579-1593.

Means, J. E., Acker, S. A., Brandon, J. F., Renslow, M., Emerson, L., Hendrix, C. J., 2000. Predicting forest stand characteristics with airborne scanning lidar. Photogrammetric Engineering & Remote Sensing 66 (11), 1367-1371.

Means, J.E., 2000. Comparison of large-footprint and small-footprint lidar systems: design, capabilities, and uses. In Proceedings: Second International Conference on Geospatial Information in Agriculture and Foretry, Lake Buena Vista, Florida, 10-12 January 2000: I-185-192.

Miller, G.L., Wagner, E. R., 1990. An optical rangefinder for autonomous robot cart Navigation. In Autonomous robot vehicles, Springer-Verlag, New York. pp.122-134

Miller, J.R., Wu, J., Boyer, M.G., Belanger, M., Hare, E.W., 1991. Season patterns in leaf reflectance red edge characteristics. International Journal of Remote Sensing 12:1509-1523.

Moffiet, T., Mengersen, K., Witte, C., King R., Denham R., 2005. Airborne laser scanning: Exploratory data analysis indicates potential variables for classification of individual trees of forest stands according to species. ISPRS Journal of Photogrammetry and Remote Sensing, 59: 289-309.

Naesset, E., 1997. Estimating timber volume of forest stands using airborne laser scanner data. Remote Sensing of Environment, 61(2): 246 – 253.

Ørka, H.O, Naesset, E., Bollandsas, O.M., 2007. Utilizing airborne laser intensity for tree species classification. ISPRS Workshop on Laser Scanning and SilviLaser 2007, Espoo, September 12-14, 2007, Finland.

Persson, A., Holmgren, J., Söderman, U., 2002. Detecting and Measuring Individual Trees Using an Airborne Laser Scanner. Photogrammetric Engineering and Remote Sensing, 68(9): 925-932.

Pfister, R. D., Kovalchik, B.L., Arno, S.F., Presby, R.C., 1977. Forest habitat types of Montana. USDA Forest Service, Intermountain Forest and Range Experiment Station, General Technical Report, INT-34. 192pp.

Popescu, S.C., Wynne, R.H., Nelson, R.F., 2003. Measuring individual tree crown diameter with lidar and assessing its influence estimating forest volume and biomass. Canadian Journal of Remote Sensing, 29(5): 564-577.

Popescu, S. C., Wynne, R. H., 2004. Seeing the trees in the forest: Using LIDAR and multispectral data fusion with local filtering and variable window size for estimating tree height. Photogrammetric Engineering and Remote Sensing, 70(5): 589−604.

Reitberger, J., Krzystek, P., Stilla, U., 2006. Analysis of full waveform lidar data for tree species classification. In Proceedings of the Photogrammetric and Computer Vision Conference, September 20-22, Bonn, Germany. 6 p.

Richardson, A. D., Reeves, J. B., Gregoire, T. G., 2003. Multivariate analyese of visible/near infrared (VIS/NIR) absorbance spectra reveal underlying spectral differences among dried, ground conifer needle sample from different growth environments. New Phytologist 161:291-301.

Rowell, E., Seielstad, C., Vierling, L., Queen, L., Shepperd, W., 2006. Using laser Altimetry-base segmentation to refine automated tree identification in managed forests of the Black Hills, South Dakota. Photogrammetric Engineering and Remote Sensing, 72(12): 1379-1388.

Rowell, E., Seielstad, C., Queen, L., 2009. Implementation and Performance of a Local-Maximum Tree Stem Identification Algorithm Using Vector-format Airborne Laser Scanner Data. Manuscript in preparation.

Schreier, H., Lougheed, J., Tucker, C., Leckie, D., 1985. Automated measurements of terrain reflection and height variations using an airborne infrared laser system. International Journal of Remote Sensing, 6(1): 101-113.

Sims, D. A, Gammon, J.A., 2003. Estimation of vegetation water content and photosynthetic tissue area from spectral reflectance: A comparison of indices based on liquid water and chlorophyll absorption features. Remote Sensing of Environment, 84(4): 526– 537.

Sollins, P., 1998. Factors influencing species composition in tropical lowland rain forest: does soil matter? Ecology 79(1):23–30.

St-Onge, B. A., 1999. Estimating individual tree heights of the boreal forest using airborne laser altimetry and digital videography. In Proceedings of the ISPRS: Mapping surface structure and topography by airborne and spaceborne lasers, November 9-11, La Jolla, USA. 6p.

Wagner, W., Ullrich, A., Ducic, V., Melzer, T., Studnicka, N., 2006. Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner. ISPRS Journal of Photogrammetry & Remote sensing, 60(2): 100-112.

Watt, P., Wilson, J., 2005. Using airborne light detection and ranging (LiDAR) to Identify and monitor the performance of plantation species mixture. Proceedings of ForestSat 2005, Borås, Sweden, pp. 51–55.

Williams, D. L., 1991. A comparison of spectral reflectance properties at the needle, branch, and canopy level for selected conifer species. Remote Sensing of Environment 35(2-3): 79-93.

CHAPTER 2

Exploratory Analysis of Airborne Laser Scanning Intensity

Abstract

An increasing number of applications of airborne laser scanning for land-cover classification rely on both structural and intensity characteristics of features. Many problems of using airborne laser scanning (ALS) intensity data have been reported and different techniques are being experimented with to minimize them. This chapter explores intensity generated by Leica ALS50 laser data and its influencing factors, including Automatic Gain Control (AGC) and range (scan angle). Two datasets are delineated from different acquisitions, one in an urban setting (MSO airport), the other in a natural setting (LEF). Simple variables are evaluated, including minimum, maximum, mean and standard deviations of intensities. Following normalization and evaluation of its effect, intensity is used to classify the two landscapes using a simple threshold method. Asphalted materials exhibit lower intensity than vegetation, and scan angle is not related to intensity. The uncorrelated intensity with the scan angle is likely due to the gain value adjustment automatically applied during acquisition. Intensity variations produced by AGC can be reduced up to 13% for asphalt materials and ~3% for natural materials following normalization. These results support other studies and emphasize the need for intensity correction prior to broader applications. Overall, intensity exhibits the potential for discriminating landscape covers. However, the thresholding method using intensity alone is not adequate for classifying the landscape and additional laser-derived variables are needed for a better result.

Introduction

In recent years, the use of airborne laser scanning (ALS) for measuring positional targets has been growing rapidly. An accurate three dimensional measurement relies on the capability of a laser system to provide x,y,z data with precise geo-control. In addition to x,y,z data, most commercial laser proprietary systems provide information on scan angle and reflectance intensity, which have been investigated for various applications, including geometric correction (Burman, 2000; Maas, 2002; Vosselman, 2002) and landscape classification (Song et al. 2002; Lutz et al. 2003; Holmgren and Persson, 2004; Hasegawa, 2006). Song et al. (2002) found that asphalt, roof, trees, and grasses were separable by applying intensity interpolation and filtering methods (not specified). They suggested that the observed 70% classification accuracy could be improved by correcting for noise derived from scan angle variation. Similarly, Lutz et al. (2003) indicated that surface classes were distinguishable as their intensity values are different, but note that the scanning geometry, including footprint size, range, and incidence angle were factors decreasing intensity values significantly. They recommended that these effects can be minimized by applying an intensity correction based on scan angle and flying height. Additionally, they note that any correction should consider the complex lidar-target interaction, as different surface types may exhibit similar intensity values that simple correction may not be adequately address the noise. Hasegawa (2006) found that intensity correction using a diffuse reflection model was difficult as shown by unchanged intensity values of new asphalt, brick and grass, even though values of old asphalt, cement, roof, and soil tile declined as scan angle increased. These studies indicate that intensity exhibits a potential for different uses, but the same studies also suggest that prior to broad applications, factors affecting intensity should be accounted for and minimized.

When the laser emits light toward a target through the scanner, it is reflected back by the target and the returned beam (echo) is captured by the receiver. Each pulse can have multiple echo/returns depending on the target and footprint coverage. Further information regarding the principal mechanism of laser scanning can be found in several publications (Balsavias, 1999; Wehr and Lohr, 1999). Because of the inability of the ALS system to generate real time accuracy of the target 3D measurement, the recorded data is synchronized during post processing with corresponding datasets generated by both differential global positioning system (DGPS) and inertial navigation system (INS) installed on the collection platform or aircraft. DGPS recalculates the position difference between the aircraft and the target while INS measures the laser orientation along three axes and the angular acceleration of the platform. This process results in georeferenced coordinates (x,y,z) of each recorded pulse adjusted to a reference coordinate system.

Meanwhile, intensity, which is a proportion of the photon numbers imposed on the detector (Balsavias, 1999) and recorded as a function of signal strength for each return, is converted into a digital number (DN). Some systems produce different measurement values to represent the level of intensity, e.g., Leica Geosystems uses 8-bit resolution (value ranges from 0-255), while the Optech system produces 32-bit resolution (value ranges from 0-8160). It is worth noting that most providers do not provide detailed insight into the proprietary pulse detection algorithms. Intensity can be a measure of maximum amplitude of the detected echo, or the integral/average of the returned signal width depending on systems. For the Leica ALS50 used in this study, intensity is adjusted to variations in slant range, flying height and system gain value using automatic gain control (AGC). Many factors affect intensity including, footprint size, range and coverage, target roughness, and reflectivity (Jelalian, 1992; Aytaç and Barshan, 2005).

A major difficulty in working with laser reflectance intensity is separation of signal from these influencing factors. Reflected laser energy is observed by a laser sensor's receiving optics as a waveform of intensity that varies with time. A sample of the waveform represents a "return" can be separated into single, first, intermediate and last return. As such, its intensity is a function of many variables including the laser optics and receiver characteristics, transmitted power, the incidence angle, atmospheric attenuation, and the path length (Baltsavias, 1999). Although intensity is not affected by shadow effects and changing illumination conditions due to the sun position with respect of the sensor, clouds, etc. (Wagner et al., 2006), the target characteristic is another factor that affects intensity significantly. A laser pulse that hits a very low reflective target may not have enough energy travel back to the receiver to be recorded as a "complete" pulse. On the other hand, target size also determines the transmitted pulse strength. It is proportional to the laser footprint diameter coverage. Additionally, the target surface roughness influences the pulse scattering direction, which ultimately affects the intensity recorded by the sensor (Steinvall, 2000).

Laser system differences are another factor influencing intensity. In general, two techniques are used to record intensity. First, intensity is recorded directly following the pulse detected by the sensor. Several commercial laser providers use this method, including Optech and Reigl. Second, intensity is adjusted to variations in slant range, flying height and system gain value using automatic gain control (AGC) and a few companies use this technique, including Leica Geosystems. The first method takes advantages of simplicity and purity of the pulse signal, in which a weak received signal is likely to have low intensity, and a strong signal exhibits high intensity. There is lack of information on the second method as to whether the AGC adjustment is satisfactory for different applications. Roth (2009) describes that for the Leica ALS50 system, the AGC adjustment takes place once in every 12-16 laser shots and it is performed automatically. A few studies note that the adjustment has a non-linear effect on intensity due to the variability of targets (Korpela, 2008). Additionally, the AGC does not always produce an acceptable adjustment, especially for a target having low signal amplitude that is dominated by noise (Adams, 2000; Wagner et al, 2004). Without question, more investigations are needed from different environments to fully understand the intensity metric.

One of several techniques used to remove or minimize noise is by correcting the intensity using a reflection model. This can be performed on a flat target, assuming that the surface target is uniform within footprint diameter and that its reflectance follows Lambert's diffuse law. The basis of this law under such conditions is that intensity is proportional to the incident angle and depends on target reflectance (Wagner et al, 2006). Experiments based on this assumption are realistic in a laboratory (Aytaç and Barshan, 2005; Jutzi and Stilla, 2006). However, it is complicated to apply it at normal state in the field. For example, Coren and Starzai (2006) noticed that a variation in incidence angle, which affects the diverging beam and atmospheric attenuation, were two primary factors influencing laser intensity. They performed a radiometric calibration to remove them using an exponential decay function (no further information). Recently, Hofle and Pfeifer

(2007) proposed two methods, data-driven and model-driven, to estimate the best parameters for range normalization using a least-squares adjustment and calculate influencing factors using a radar system model respectively. They used airborne laser datasets acquired at different heights and found that both methods were able to reduce intensity variation and geometric offsets between flight strips. When lidar data is collected from significantly different elevation ranges, such as in mountainous landscapes, the effect of range differences also need to be accounted for (Balsavias 1999). Luzum et al. (2004) considered the signal loss due path length variation, and it was necessary to compensate by multiplying the raw intensity by the squared range, divided by a user-defined standard range. Clearly, more studies are needed to explore intensity characteristics, especially using datasets generated by a system operating with AGC.

The primary goals of this chapter are to understand intensity and how AGC affects the intensity, and to evaluate the effectiveness of intensity normalization for range (scan angle) and AGC. Objectives are: (1) to identify/quantify intensity characteristics and influencing factors, including range (scan angle) (2) to normalize intensity and to quantify the quality of normalizations, and (3) to use normalized data to classify landcover in both urban and natural settings including homogeneous stands of Douglas-fir, ponderosa pine, lodgepole pine and western larch.

Materials and Methods

Study sites

The study area is located in and around the Missoula Valley, in western Montana (46° 55' N and 114° 05 W, elev. 973 m) consisting of two sites. The first site is the Missoula International airport (MSO) and the second site is The University of Montana's Lubrecht Experimental Forest (LEF), 40 km northeast of Missoula (approximately N 46° 53' W 113° 27', with elevations from 1160 to 1930 m). Figure 1 shows an overview of the study sites.

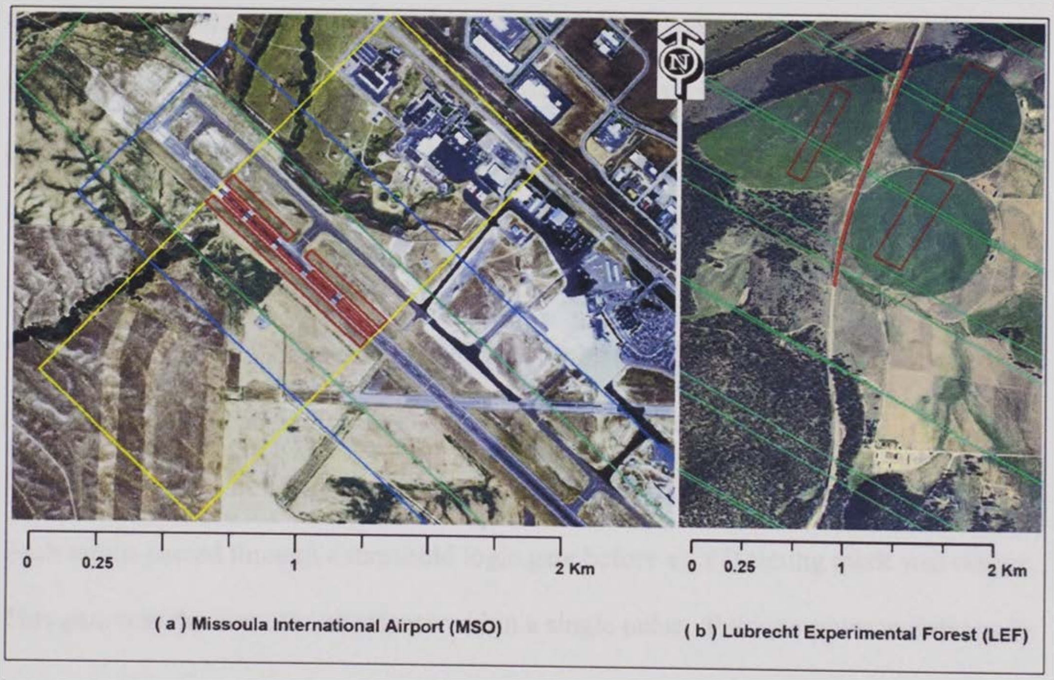


Figure 1. Overview of study area with flight strips and samples (red).

Lidar dataset

ALS data were acquired on July 27 and August 17, 2005, and June 26, 2006 using Leica Geosystems Proprietary ALS50. At MSO, four flight strips were obtained, three of them are shown in Figure 1. Strip 1 was acquired at 1250 m above mean terrain elevation (AMTE) perpendicular with the runway (Southwest-Northwest) while the Strip 2 and Strip 3 were scanned, at 750 m and 1250 m respectively, in the same direction as the runway (oriented to Southeast-Northwest). At LEF, 54 flight strips were obtained at the mean flying height of 1829 meters AMTE, mean airspeed was 72 m/s, scan rate was 35KHz, and scan angle was $\pm$35 degrees. These parameters resulted in an average return density of ~1 return per 2.29 m^2 on the ground (multiple returns and sidelap result in up to 5 returns/m^2 depending on site), while a laser beam divergence of 1 mrad (1064 nm) combined with flying height resulted in an average footprint size of ~ 1 m^2. The sidelap was 50% to ensure complete coverage of the study area. A maximum of four returns for each pulse were recorded depending on vegetation height and pulse return energy following first hits. A Constant Fraction Discriminator (CFD) was used to produce a timing mark at the half-max amplitude on the leading edge of a pulse (Roth, 2007). Intensity was measured at the peak of the return pulse and digitized as an 8-bit value. Each return passed through a threshold logic gate before a CFD timing mark was output. This gate was the same for all returns within a single pulse. Pulse-to-pulse variability in laser output is reported at less than 5%. A summary of acquisition parameters for this study is shown in Table 1. The signal to noise ratio (SNR) for the acquisition was greater than 10:1 for 10% diffuse reflectivity targets (Roth, 2007). The vertical accuracy is reported at 0.15 m and horizontal accuracy is 0.25 m. Following pre-processing (roll, pitch, and yaw corrections), above ground laser return points were separated from ground 'bare earth' points using the triangulated irregular network (TIN) densification method available in the TerraScan software suite (Terrasolid, 2004). A Digital Elevation Model (DEM) was created using Inverse Distance Weighted (IDW) interpolation at 1 m resolution and used to calculate the Canopy Height Model (CHM) using a spot elevation method. This technique computed the canopy height of each point by subtracting the DEM height from the CHM.

The Leica ALS50 used for this study operated with Automatic Gain Control (AGC) to record target intensity, which automatically adjusts the captured raw intensity values for variations in slant range, flying height, and system gain (Leica Geosystems, 2008). This adjustment may have a nonlinear effect on the output target intensity due to reflectance variability. For the purpose of this study, intensity correction was carried out to minimize such effect by normalizing the raw intensity with squared range divided by squared average elevation, multiplied by AGC value (Korpela et al. 2009).

Table 1. ALS system characteristics.

Specification	Missoula International Airport (MSO)	Lubrecht Experimental Forest (LEF)
Date of Acquisition	July 27, 2005	July 27 and August 17, 2005
ALS system	Leica Geosystems ALS50	Leica Geosystems ALS50
Ave. Flight Height above Surface	750 and 1250 m	1928 m
Ave. Flight Speed	72.02 m/sec	72.02 m/sec
Number of Strips	4	54
Scan frequency	35 Hz	35 Hz
Laser Pulse Frequency	43 KHz	43 KHz
Laser Power	4 Watts	4 Watts
Signal to Noise for diffuse targets	-10% diffuse = 30:1 - 5% diffuse = 20:1	-10% diffuse = 12:1 - 5% diffuse = 7:1
Attenuation setting	50% and None	50%
Scan Angle	± 35°	± 35°
Sidelap	-	50%
Average Swath Width	473 m and 788 m	1153 m
Average Return Density	1.28 per m^2 and 0.77 per m^2	0.44 m^2
Average Area / Return	0.78 m^2 and 1.30 m^2	2.29 m^2

Sample delineation

In order to evaluate intensity characteristics, different large homogeneous cover types are located in both MSO and LEF. Samples were delineated from both sites using the following steps. First, an approximately homogeneous target was pre-selected on each site using the color digital aerial photo of National Agricultural Imagery Program (NAIP) of U.S. Farm Service Agency acquired in summer 2005 (USFSA, 2007). Homogeneity was subjectively determined by identifying areas with consistent tone, texture and color in the imagery. Second, a rectangular polygon was placed on the target (Figure 1). At MSO, two samples were selected: (1) runway, and (2) grasses around the runway. The runway sample was delineated manually to exclude white marks and overused take off/landing areas marked by aircraft tires (usually darker/black). The grasses were sampled from a cleared area adjacent to and on both sides of the runway (called the runway strip). For LEF, two samples were delineated: (1) alfalfa, delineated from three locations and assumed to have similar height variation, and (2) highway 200, located on a relatively flat area running orthogonally to the flight strips. The white marks on the highway were not removed due to unclear appearance on the NAIP image.

Bare earth laser returns containing x,y,z, intensity and angle values were extracted from inside polygons of each sample. In order to analyze the effect of scan angle (range), returns were grouped into scan angle classes at interval of 1° (from -14° to 14°) on both sides of the nadir. The recorded intensity has range of values from 0-255 (8 bit). The zero value is an underflow that results from the ALS algorithm's attempt to represent a too small number while the value of 255 is the maximum value in the 8-bit format. Both values represented saturation and were removed from the samples to minimize a noise (<0.47% of total returns).

Intensity normalization

Intensity normalization was carried out on an individual point/return basis with the goal of reducing the intensity variation occurring in the samples. Korpela (2008) proposed a calibration model to achieve a similar goal by calculating the raw intensity with squared range divided by squared average elevation, multiplied by AGC value using the following equation:

$$Int_{norm} = Int_{Raw}\left(\frac{R}{R_s}\right)^a + (b.Int_{Raw}).(127.5 - AGC) \quad (1)$$

where Int_{norm} is the normalized intensity value, R is the range, R_s is the standard range, AGC is the system gain value, and a and b are normalization parameters. Baltsavias

(1999) suggests that return intensity of a homogeneous target is inversely proportional to the second power of the range. Therefore, the range value of the parameters *a* was evaluated from 2.00-3.00 while the parameter *b* was tested from 0.004-0.005. The constant value of 127.5 was adjusted to sample characteristics, including the minimum, maximum and mean value of the AGC. The normalization was performed by applying the equation 1 with some possible combinations of both parameters and the constant value. The greatest difference in the intensity variation between before and after normalizations was used to indicate the efficacy of the normalization and to estimate the AGC effects on intensity. Overall 24 equations were examined. It is worth noting that there was very little difference in results between equations. The equation that reduces variability the most was with a=2.0, b=0.05, and constant value of 127.5. Figure 2 shows the result of normalization using Inverse Distance Weighting (1 meter resolution) with sample polygons.

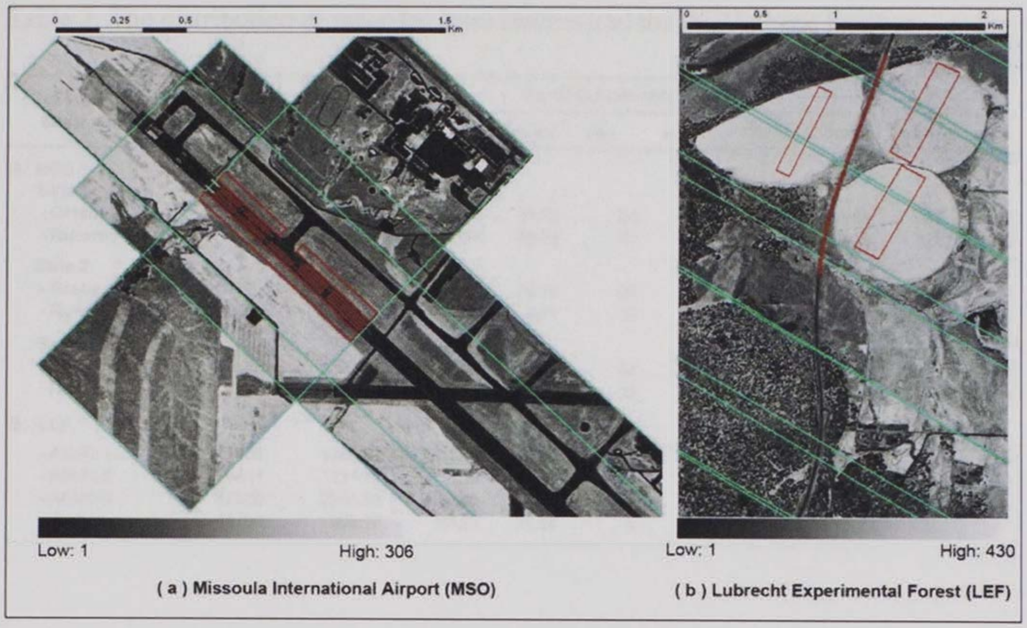


Figure 2. Overview of study area with flight strips and samples (red) following normalization.

Basic statistical analysis and landscape classification

Intensity variables were calculated from samples of both study sites, including means, maximums and minimums, standard deviations, and coefficient of variation (CV). The mean, minimum and maximum of the automatic gain control (AGC) values were also calculated. Table 2 shows distributions of sample returns and recorded intensity values.

Table 2. The distribution of recorded laser returns and the AGC characteristics.

Site and cover type	Total returns	Average returns/angle	Recorded intensity					AGC value		
			Mean	Std.Dev	Min.	Max.	CV	Mean	Min.	Max.
A. MSO										
Strip 1										
- Grass	42609	236.44	161.70	21.09	54	254	0.13	125.62	122	133
- Runway	9923	966.27	119.84	30.26	47	196	0.25	139.49	125	150
Strip 2										
- Grass	47452	-	158.18	25.11	53	254	0.16	128.90	126	135
- Runway	10639	-	51.52	8.11	19	89	0.16	129.22	126	135
Strip 3										
- Grass	30434	-	160.85	21.34	95	254	0.13	124.61	122	128
- Runway	6890	-	52.17	6.86	26	82	0.13	124.81	123	129
B. LEF										
- Alfalfa 1	37148	1280.97	221.80	11.36	18	254	0.05	121.63	120	124
- Alfalfa 2	55441	1911.76	215.38	14.80	91	254	0.07	122.13	120	124
- Alfalfa 3	67539	2701.56	215.89	11.56	55	254	0.05	121.48	121	124
- Highway	10404	395.76	83.62	31.84	6	254	0.38	123.33	121	128

For the purpose of exploring the usefulness of intensity for landscape analysis, a simple classification was performed on both MSO and LEF using threshold values (minimum and maximum) selected manually. The classification was carried out in grid format (3x3m resolution) and the nearest neighborhood was used for converting points to grids. The MSO was classified into three primary cover types: (1) asphalt, (2) grass, and (3) bare soil; while LEF was grouped into: (1) vegetation, (2) bare soil/grass, and (3) alfalfa. General statistic values were calculated for each cover type, including mean, minimum and maximum, standard deviation values. Separate samples from each site were delineated with the aid of the NAIP image and the stand polygon database at MSO and LEF respectively. Samples in MSO including trees, building roof, and mix grasses/shrubs were selected and delineated manually while the stands at LEF were selected using the criteria determined by the stand database (Waterman, 2000). Five stands approximately representing the range of intensity characteristics across LEF were chosen, including ponderosa pine, mixed Douglas-fir, mixed ponderosa pine, and mixed western larch. It should be noted that the name of the stand indicates the prime/dominant species within each class. Similar basic statistical values were calculated and vegetation samples (shrubs and trees) were computed from average values of the bare earth and canopy height model (CHM) returns. The characteristics of cover types are presented in Table 3.

Results

Intensity characteristics

The recorded intensity characteristics of cover types in MSO and LEF are shown in Table 2. In general, the intensity of green vegetation (grass and alfalfa) is higher than asphalts (runway and highway). At MSO, the intensity of the runway delineated from Strip 1 (the flying direction was perpendicular with the runway) is higher than intensity from Strips 2 and Strip 3 (the flying direction is parallel with the runway). Strip 1 is also represented by higher intensity variation (coefficient of variation/CV) when compared to other samples, including grasses. The mean value of automatic gain control (AGC) is within a close range between samples, except the runway from Strip 1, which is higher than others. At LEF, intensities and variations of all three alfalfa samples are similar while the intensity variation of the highway is slightly higher. The AGC of both cover types are comparatively similar to each other. Overall, asphalt appears to exhibit a higher intensity variation (subsequent to the reduction following the normalization) than grass regardless of sample sites.

Figure 3 illustrates the trend of the recorded intensity with scan angle. At Missoula International Airport (MSO), the intensity of grass is variable relative to the other classes, but it does not change with the scan angle. However, the intensity of the runway shows more variation between two sides of the nadir ($\alpha<0^\circ$ is the left side and $\alpha>0^\circ$ is the right side). Return intensity from the left side ($\alpha<0^\circ$) is generally lower than the right side ($\alpha>0^\circ$). The cause of this observation is unknown. Nonetheless, it does not signify a decrease with an increase of the scan angle. At Lubrecht Experimental Forest (LEF), both alfalfa and highway intensities are consistent across scan angles as would be expected in absence of a scan angle effect. The intensities on both sides of nadir are characterized by relatively similar values. Overall, all samples indicate that the recorded intensity does not exhibit a trend with scan angle.

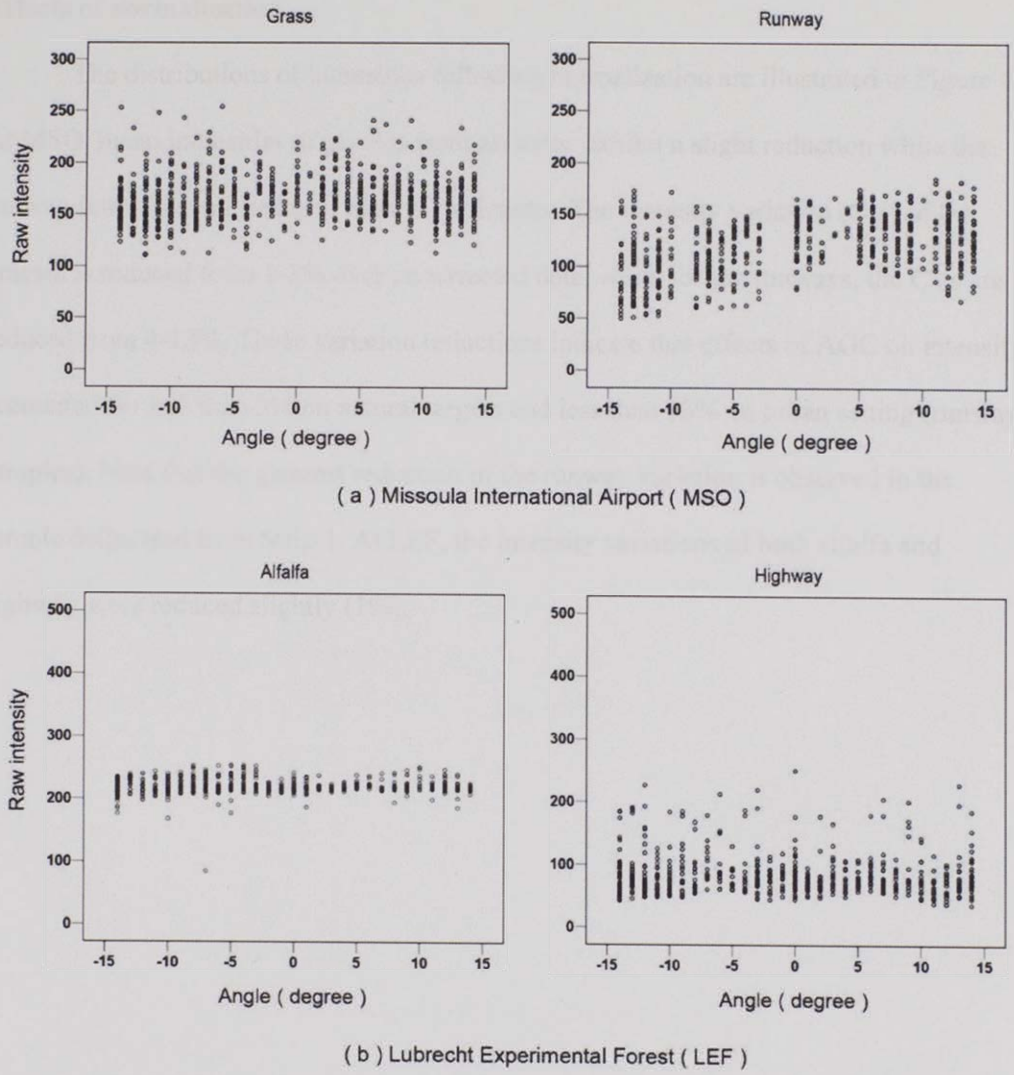
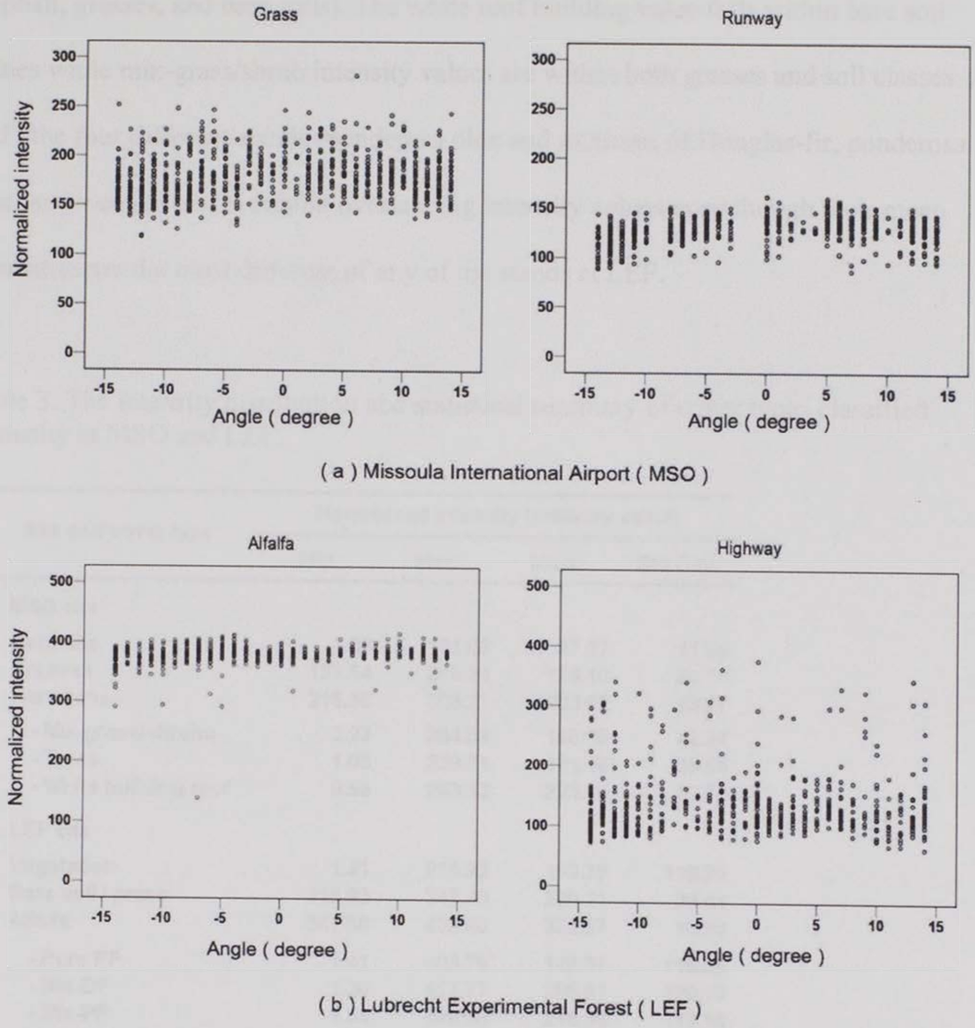


Figure 3. The recorded intensities of return samples from MSO and LEF (selected randomly from 800 returns for each cover type)

Effects of normalization

The distributions of intensities following normalization are illustrated in Figure 4. At MSO, mean intensities of grasses from all strips exhibit a slight reduction while the runway intensities are reduced more significantly. The intensity variation (CV) of the grasses is reduced from 1-3% over uncorrected data, while for the runways, the CVs are reduced from 4-13%. These variation reductions indicate that effects of AGC on intensity accounted for less than 3% on natural targets and less than 13% on urban setting (runway samples). Note that the greatest reduction of the runway variation is observed in the sample delineated from Strip 1. At LEF, the intensity variations of both alfalfa and highway were reduced slightly (1%).



(a) Missoula International Airport (MSO)

(b) Lubrecht Experimental Forest (LEF)

Figure 4. The normalized intensities of return samples from MSO and LEF (selected randomly from 800 returns for each cover type).

Landscape classification

Table 3 shows threshold values used for the classification and the general statistics of cover types at both sites. At MSO, the mean intensity of asphalt is the lowest while bare soil is the highest. The range of intensity values of other cover types (mix-grass/shrubs, trees, and white roof) falls within the threshold values of main cover types (asphalt, grasses, and bare soils). The white roof building value falls within bare soil values while mix-grass/shrub intensity values are within both grasses and soil classes. At LEF, the four different stands (ponderosa pine and mixtures of Douglas-fir, ponderosa pine, and western larch) exhibit overlapping intensity values even though their mean intensities are the most different of any of the stands at LEF.

Table 3. The intensity distribution and statistical summary of cover types classified manually at MSO and LEF.

Site and cover type	Normalized Intensity (arbitrary value)			
	Min.	Max.	Mean	Std.Dev
A. MSO site				
Asphalts	1.03	131.63	47.97	11.58
Grasses	131.64	215.34	168.10	20.14
Bare soils	215.35	306.21	223.92	13.27
- Mix-grass/shrubs	3.23	284.84	198.98	42.94
- Trees	1.06	239.71	121.16	66.55
- White building roof	9.56	283.13	223.10	30.58
B. LEF site				
Vegetation	1.41	215.92	199.38	119.79
Bare soil / grass	215.93	347.49	289.31	36.44
Alfalfa	347.50	430.40	372.57	16.33
- Pure PP	1.41	403.79	148.37	116.25
- Mix-DF	1.38	417.77	206.81	120.13
- Mix-PP	1.34	399.06	216.34	113.10
- Mix-WL	1.23	390.21	198.02	122.80

Results of the simple classification and some misclassified cover types are shown in Figure 5. At MSO, clearly, asphalt materials are significantly distinguished from grasses and bare soils. Incorrectly classified covers are also observable, including a cluster of trees grouped into asphalt and mix-grass/shrubs, and a building roof categorized as bare soils. At LEF, most alfalfas are obviously separated from high vegetation/trees and bare soil/grass. At the lower right of the subset, a small portion of bare soil/grass is mistakenly classified into alfalfa while the highway and gravel roads are classified into vegetation. However, all stands are not clearly distinguished from one another. The mix Douglas-fir stand appears similar to the other three stands.

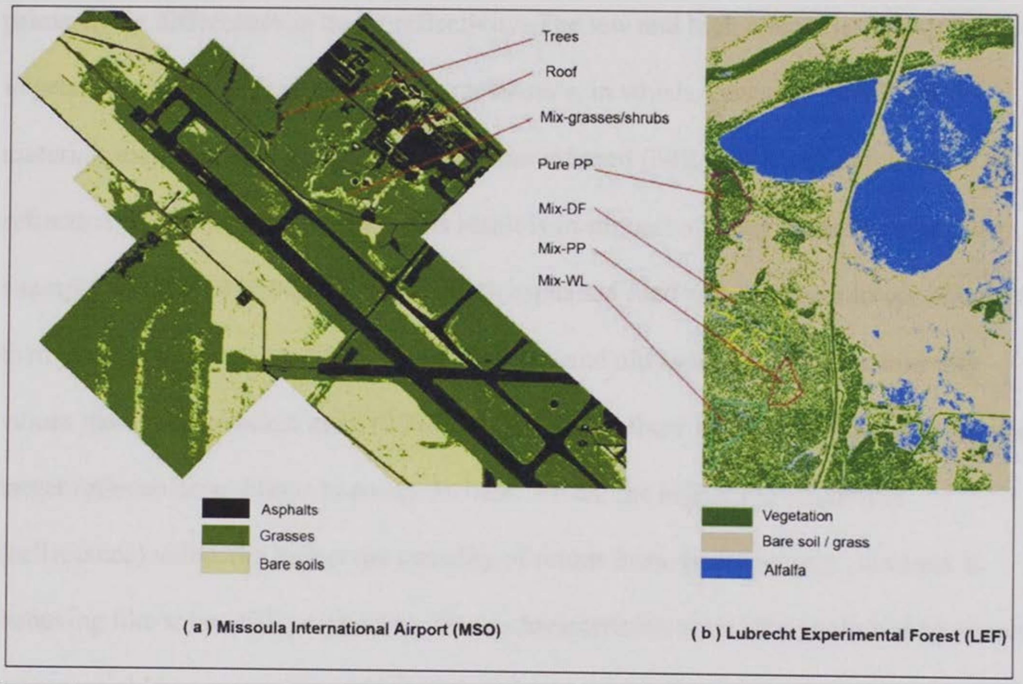


Figure 5. A simple landscape classification.

Discussion

The basic property of laser intensity is well documented, but some factors affecting it are still difficult to quantify. My goal is to explore the intensity characteristics and the effectiveness for possible landscape classification. The analyses are limited to the influencing factors of scan angle (range) and automatic gain control (AGC) and use basic statistical metrics. Results of two datasets (Missoula international airport and Lubrecht Experimental Forest) provide some concepts about intensity and those two factors.

Generally, the effect of AGC on intensity is visible as the intensity does not change significantly with either the scan angle. The cover types exhibit differences in intensity and a general landscape classification can be performed using a simple method.

The difference in intensity between cover types (Table 2) is likely explained primarily by differences in target reflectivity. The low and high intensities of asphalt and vegetation are due to their spectral characteristics, in which, generally, asphalted materials exhibit a high absorption in the near infrared (NIR) while vegetation has a high reflection within this wavelength. This result is in support of other investigations. For example, Song et al. (2002) noted that both asphalted road and roof have lower intensity than grass and Hasegawa (2006) found that new and old asphalts show less intensity values than grass. Ahokas et al. (2008) suggested that there is strong relationship between target reflectance and laser intensity. In other words, the higher the brightness (reflectance) value, the higher the intensity of return from the target (e.g., the laser is behaving like solar (NIR) radiation). These characteristics were also published by several commercial laser companies, which included other reflectivity data of various materials at different wavelengths (Riegl, 2009; Optech, 2009).

With regard to the observed variations, the results indicating that intensity varies within cover types are anticipated, as the composition of each cover type is not truly homogeneous. However the between samples homogeneity assumption appears sufficient given the slight differences in variation. There is no observable difference in intensity and gain values among cover types delineated by the same flight directions. The differences observed from the flying strip perpendicular with the runway (Strip 1), especially asphalt, are probably due to a non-linear effect of the AGC adjustment as well as the cover composition. Although it is difficult to quantify such effect due to the fact that the intensity algorithm and information regarding the effectiveness of the AGC are not available, the non-linear effect is likely generated by a scanning density and intensity variability of cover types for a particular scanning period. Although Roth (2009) suggests that for the ALS50 system, the AGC adjustment automatically takes place once in every 12-16 laser shots, detailed observations performed separately on both MSO and LEF samples do not indicate intensity differences. However, it is observed that such adjustment may result in a pattern where similar intensity and AGC values are used to adjust different cover types within those 12-16 laser shots. For example, the runway intensity values delineated from Strip 1 may be adjusted based an AGC values of grasses within that area due to the geometry of the flight axis, runway, and taxiway. It is worth noting that the results for Strip 1 are not intuitive and may point to an error with the AGC or to a lack of detailed understanding of how the AGC algorithm is implemented.

The absence of a scan angle effect (Figure 3) may be also explained by the AGC adjustment, in addition to the narrow range of the scan angle classes (±14°). As discussed above, intensity is indeed already adjusted to the scan angle (slant range) by the AGC and it is likely the primary reason for such consistency. On the other hand, effects of a narrow scan angle range in influencing intensity may not be as significant as the AGC effect, given the fact that the adjustment is carried out automatically. It is acknowledged that to quantify effects over narrow angles remains challenging due to lack of data regarding methodology/algorithms used for intensity and AGC calculations. In addition, there are no other datasets acquired at both MSO and LEF using a different ALS system for a comparison. However, it is worth noting that correlations between intensity and scan angle vary regardless of systems used for the acquisition. For example, using different ALS systems (without AGC) for evaluating intensity characteristics on different targets (gravels, grasses, and tarps), Kukko et al. (2008) found that the intensity variability was low within the incident angle <30° and the effect of the angle on intensity was not significant. However, they suggested that such effect is stronger on high reflectance target and more studies were needed to explain other factors, such as surface roughness and laser footprint. In contrast, Coren and Sterzai (2006) and Hofle and Pfeifer (2007), also using non-AGC ALS systems noted that intensity was significantly affected by scan angle. Clearly, characterizing laser intensity is difficult. It is worth mentioning that a separate calculation conducted by the author using samples from different elevations also indicates that intensity is not affected by the range difference.

The difference in reduction values of intensity variations indicate that each cover type exhibits different intensity characteristics, which are likely affected by both range and the AGC adjustments mentioned above. Although a large change following normalization is only observed on the asphalted material (runway); (Figure 4), it is suggested that intensity should be corrected prior to applying it to landscape classification. Additionally, different normalization methods may result in different variation reductions, depending on the ALS system, targets, and acquisition setting (e.g., flying height, strips, and pulse repetition frequency). For example, using Leica ALS50 II (the pulse repetition frequency (PRF) is almost three times higher that the one used in this study), Korpela (2008) found that the AGC and range effects on intensity were 27% and 20% respectively. They also noted that the overall accuracy of an LDA application for discriminating lichen from soil and other understory increased 4% following normalization.

While the results of the simple classification presented in this chapter clearly show that laser intensity data can be used to discriminate between general landscape classes, it produces many misclassifications of features that may be of interest to future studies. The misclassifications occurring on both the urban and forested sites are not unanticipated given that the method used is based on entirely static threshold values. The results suggest that intensities of the observed cover types are within the range values of each other, and a more robust methodology with additional data is needed to have a chance of resolving ambiguous classifications of these features. For example, the misclassification of several trees into the asphalt class in MSO is likely due to both structural and species characteristics of these trees (mostly ponderosa pine) that reflect lower intensity value than average values (Table 3). On the other hand, the incorrectly classified mix-grasses/shrubs into bare soil class is not uncommon as both green, low vegetation and dry soil exhibit similar high intensity, depending on the several influencing factors mentioned above. For example, Hasegawa (2006) found that the range of tree intensities was within the range of soil, and standard deviations of intensities of both covers overlapped. He also suggested that there was a clear difference in intensity between grass and soil. Therefore, it is proposed that while the landscape stratification at a general level is possibly using ALS datasets, a simple methodology, such as the one used in this study is not sufficient to produce the more detailed landscape classes required of many forestry applications.

The results of classification in LEF emphasize the need of more rigorous methodologies for discriminating specific forest cover types. Overlapping intensity values are often difficult to separate, especially if these values are due to similarities in physical characteristics. For example, in the field, part of bare soil/grass misclassified into alfalfa is dominated by grass (Figure 5), which is mostly green in color reflecting high intensity in the NIR similar to alfalfa. The inclusion of bare soil/grass inside forest stands may be also due to such similarity. It is important to recognize that the threshold values used for this classification are generated from both bare earth and canopy returns, which affect intensity distribution of those stands. A separate calculation using only vegetation returns shown in Table 3 also indicate that intensities of the four stands are within the range of each other. The result is expected, noting that intensity characteristics of trees can vary even between similar species depending on their structures and reflectivity (Holmgren and Persson, 2004; Moffiet et al., 2005). However, the apparent difference in mean intensities between the four stands used in this calculation suggests that species classification may be possible using airborne laser scanning intensity datasets, in conjunction with structural variables. An improved technique using more variables derived from ALS system, such as canopy height, spatial distribution, and separated return types will likely classification. This concept is exploited in a following chapter (Chapter 4).

Conclusion

In this chapter, two airborne laser scanning datasets acquired with Automatic Gain Control are explored to identify factors that affect intensity characteristics. A simple classification is performed to determine whether intensity has the potential to allow discrimination between land covers. There is a general lack of detailed understanding of how intensity converted into digital number (DN) and how AGC algorithm is implemented due to proprietary behavior on the vendor's part. Therefore, while it is important to acknowledge such limitations, I also recognize that many restrictions are applied in this exploration, such as the generality of statistical metrics used and the homogeneity assumption for analyzing the data. Indeed, the primary goals of this chapter are to gain knowledge about intensity characteristics and how AGC affects the intensity, and to assess the effectiveness of intensity normalization for range (scan angle) and AGC. In general, the recorded intensity is different between cover types. The NIR reflectance value of each cover type is quasi-intuitively linked to intensity, in which asphalt reflects lower intensity than vegetation. The greater intensity variation on asphalt than vegetation is likely due to the effect of the automatic gain control adjustment. The observed slight change intensity with scan angle also may be due to such adjustment. Following normalization, intensity variations are reduced by up to 13% and 3% for asphalt and vegetation respectively. In general, intensity indicates the potential for landscape classification. However, it is not likely that a thresholding method using intensity alone is sufficient to discriminate species at stand levels. More laser-derived variables and a better methodology are needed to classify land covers into meaningful categories. For example, in order to classify landscape having high variability in the structure and composition like LEF, two stages classifications are performed, including developing Linear Discriminant Analysis based species data and performing Maximum Likelihood Classification method. The details of this methodology are described in Chapter 4.

References:

Adams, M.D., 2000. Lidar design, use and calibration concepts for correct environmental detection. IEEE Transactions on Robotics and Automation,16(6), 753-761.

Ahokas, E., Kaasalainen, S., Hyyppä, J., Suomalainen, J., 2006. Calibration of the Optech ALTM 3100 laser scanner intensity data using brightness targets. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 1).

Aytaç T, Barshan, B., 2005. Surface differentiation by parametric modeling of infrared intensity scans, Optical Engineering, 44(6),1-9.

Baltsavias, E.P., 1999. Airborne laser scanning: basic relations and formulas. ISPRS Journal of Photogrammetry and Remote Sensing 54 (2–3), 199–214.

Burman, H., 2000. Adjustment of laserscanner data for correction of orientation errors. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 33 (Part B3/1), 125–132.

Coren, F., Sterzai, P., 2006. Radiometric correction in laser scanning. International Journal of Remote Sensing 27 (15-16), 3097-3104.

Hasegawa, H., 2006. Evaluations of LIDAR reflectance amplitude sensitivity towards land cover conditions. Bulletin of the Geographical Survey Institute 53(3), 43-50.

Hofle, B., Pfeifer, N., 2007. Correction of laser scanning intensity data: Data and model-driven approaches. ISPRS Journal of Photogrammetry & Remote Sensing 62 (2007), 415–433.

Holmgren, J., Persson, A., 2004. Identifying species of individual trees using airborne laser scanning. Remote Sensing of Environment 90 (4), 415-423.

Jelalian, A.V., 1992. Laser Radar Systems. Artech House, Boston.

Jutzi, B., Stilla, U., 2006. Range determination with waveform recording laser systems using a Wiener Filter. ISPRS Journal of Photogrammetry and Remote Sensing 61(2), 95–107.

Korpela, I.S., 2008. Mapping of understory lichens with airborne discrete-return LiDAR data. Remote Sensing of Environment 112 (10), 3891–3897.

Kukko, A., Kaasalainen, S., Litkey, P., 2008. Effect of incidence angle on laser scanner intensity and surface data. Applied Optics, 47(7).

Lutz, E., Geist, T., Stötter, J., 2003. Investigations of airborne laser scanning signal intensity on glacial surfaces: Utilizing comprehensive laser geometry modeling and orthophoto surface modeling (A case study: Svartisheibreen, Norway). International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 34 (Part 3/W13), 143–148.

Luzum, B., Starek, M., Slatton, K.C., 2004. Normalizing ALSM intensities. GEM center report No. Rep. 2004-07-001. University of Florida. 8pp.

Maas, H.G., 2002. Methods for measuring height and planimetry discrepancies in airborne laserscanner data. Photogrammetric Engineering and Remote Sensing, 68 (9), 933–940.

Optech, 2009. Optech: ALTM Terrestrial Survey. http://www.optech.ca/altmhow.htm (accessed on April 16, 2009).

Riegl, 2009. Riegl Laser Measurement System. http://www.riegl.com/products/airborne-scanning/produktdetail/product/scanner/6/ (accessed on April 16, 2009).

Song, J.H., Han, S.H., Yu, K., Kim, Y.I., 2002. Assessing the possibility of land-cover classification using lidar intensity data. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 34 (Part 3B), 259–262.

Steinvall, O., 2000. Effects of target shape and reflection on laser radar cross sections. Applied Optics, 39 (24): 4381-4391.

USFSA, 2007. U.S Farm Service Agency: National Agricultural Imagery Program. http://nris.mt.gov/nsdi/orthophotos/naip_2005.asp (accessed on March 22, 2007).

Vosselman, G., 2002. On the estimation of planimetric offsets in laser altimetry data. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 34 (Part 3A), 375–380.

Wagner, W., Ullrich, A., Melzer, T., Briese, C., Kraus, K., 2004. From single-pulse to full-waveform airborne laser scanners: potential and practical challenges. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 35 (Part B3), 201–206.

Wagner, W., Ullrich, A., Ducic, V., Melzer, T., Studnicka, N., 2006. Gaussian decomposition and calibration of a novel small-footprint full-waveform digitizing airborne laser scanner. ISPRS Journal of Photogrammetry & Remote Sensing, 60, 100-112.

Waterman, J.R., 2000. Building a Lubrecht forest inventory internet mapping service and volume comparisons. The Univeristy of Montana. Missoula, MT.

Wehr, A., Lohr, U., 1999. Airborne laser scanning: an introduction and overview. ISPRS Journal of Photogrammetry and Remote Sensing, 54(2-3): 68–82.

CHAPTER 3

Tree Species Identification in Mixed-Coniferous Forest Using Airborne Laser Scanning

Abstract

This study tests the capacity of relatively low density (< 1 return/m^2) airborne laser scanner data for discriminating between Douglas-fir, western larch, ponderosa pine, and lodgepole pine in a western North American montane forest and it evaluates the combined effect of intensity, height, and return-type metrics for classifying tree species. Collectively, Exploratory Data Analysis, Pearson Correlation, ANOVA, and Linear Discriminant Analysis show that structural and intensity characteristics generated from LIDAR data are useful for classifying species at dominant and individual tree levels in multi-aged, mixed conifer forests. Proportions of return-types and mean intensities are significantly different between species (*p-value* < 0.001) for plot-level dominant species and individual trees. Classification accuracies based on single variables range from 49-61% at the dominant species level and 37-52% for individual trees. The accuracy can be improved to 95% and 68% respectively by using multiple variables. The inclusion of proportion of return-type greatly improves classification accuracy at the dominant species level, but not for individual trees, while canopy height improves accuracy at both levels. Overall differences in intensity and return-type between species largely reflect variations in the physical structure of trees and stands. These results are consistent with the findings of others and point to airborne laser scanning as a useful source of data for species classification. However, there are still many knowledge gaps that prevent accurate mapping of species using ALS data alone, particularly with relatively sparse datasets like

the one used in this study. Further investigations using other datasets in different forest types will likely result in improvements to species identification and mapping for some time to come.

Keywords: North America; conifer; laser scanning; intensity; tree species.

Introduction

Airborne laser scanning (ALS) has recently become a key technology for generating measurements to support forest inventories (Means et al. 2000; Gobakken and Naesset 2004; Maas et al. 2008). The ability of laser scanning to provide direct estimates of canopy structural features such as height, canopy closure, and stem density is highly advantageous in the inventory data collection. However, one important limitation of ALS-derived inventories is the difficulty in identifying tree species. Consequently, foresters continue to rely on passive remote sensing classifications, field surveys, and a priori knowledge of vegetation distributions to generate species data (Cochrane, 2000; Asner et al., 2002). Because landscape-level acquisitions of traditional remote sensing (imagery) and ALS data combined with field measurement are considered expensive, an alternative scheme is to optimize the usefulness of laser scanning data, by expanding its ability to provide additional information content such as tree species. Several researchers have speculated that species discrimination is possible using laser data directly (Holmgren and Persson, 2004; Moffiet et al., 2005; Brandtberg, 2007; Ørka et al., 2007), in part because most modern laser scanning systems record the intensity of individual returns.

The most recent studies conducted in Scandinavian, Australian and Scottish forests respectively, have shown that intensity is useful to distinguish between different tree species, particularly when used in conjunction with structural variables such as return type and proportions of heights and canopy hits (Holmgren and Persson, 2004; Moffiet et al., 2005; Donoghue et al., 2007; Ørka et al., 2007). These analyses were mostly exploratory and conducted to discriminate between conifer and deciduous tree species using high density data (3-10 returns/m^2). For example, Holmgren and Persson (2004) applied linear and quadratic discriminant analyses to differentiate between individual Scots pine, Norway spruce and deciduous trees in a Scandinavian boreal forest using both lidar intensity data and tree crown shape. The proportion of returns located above crown base height, the standard deviation of intensity and height, and the proportion of first returns were helpful for species identification. Moffiet et al. (2005) used vegetation permeability (vegetation points/all points) and singular vegetation returns to distinguish between White cypress pine and Poplar box at the dominant species level (the species having the greatest proportion of canopy area for each subplot by stand) and found that the proportion of singular returns contributed most of the discriminatory power (overall classification accuracy ~77%). A clear distinction between the two species was not apparent at the individual tree scale. Donoghue et al. (2007) compared normalized intensity for spruce and pine trees in Galloway forest of Scotland using one-way ANOVA to show that discrimination was possible using mean intensity (*p-value* < 0.05). The most recent study, conducted by Ørka et al. (2007), identified spruce, birch and aspen in Norway with Principal Component Analysis (PCA) and linear discriminant analysis (LDA). These authors used a combination of five components (mean intensities of first and second returns, standard deviations of first, second and single returns) to generate a classification with overall accuracy of 74%.

Each of the aforementioned studies utilized relatively high-density datasets that are generally obtained using slow, low-flying, rotor-wing aircraft. In western North American forests, ALS systems are typically flown aboard fixed-wing aircraft at higher altitudes, usually resulting in data densities more than an order of magnitude smaller in size than those obtained from typical helicopter acquisitions. Additionally, complex terrain, large environmental gradients, and mixed species composition/ forest structure complicate classification of vegetation properties. In western Montana, U.S.A., the location of our study area, forests occupy 7.5 million hectares, and are mostly dominated by multi-aged ponderosa pine (*Pinus ponderosa*), Douglas-fir (*Pseudotsuga menziesii*), western larch (*Larix occidentalis*) and even-aged lodgepole pine (*Pinus contorta*) (Arno, 1979). Correctly classifying these four conifer species is a pre-requisite of forest inventory in the region, and if ALS is to be used, the size of the landscape dictates the use of relatively sparse laser scanner data.

This study tests the capacity of relatively low density (< 1 return/m^2) ALS structure and intensity data for discriminating between the four previously cited tree species in a montane mixed conifer forest at dominant species and individual tree scales. Exploratory Data Analysis (EDA), Pearson correlation, one-way ANOVA, and Linear Discriminant Analysis (LDA) are used at the levels of plot-level dominant species and individual tree to characterize (1) differences/similarities in the proportion of vegetation return types as a function of tree species, (2) the relationships between return types and intensities, and (3) many of the intensity variables that are most useful for classifying species. This work corroborates the findings of others studying tree species discrimination in Europe and Australia, but differs from previous research in that all combinations of discriminating variables are examined.

Materials

The study site is The University of Montana's Lubrecht Experimental Forest (LEF) in the Blackfoot River drainage, 40 km northeast of Missoula, Montana (approximately N 46 ° 53' W 113 ° 27', with elevations from 1160 to 1930 m) as shown in Figure 1. Approximately 45% of the area has a slope gradient over 30%, with the steepest gradients exceeding 90% (Nimlos, 1986). LEF covers 11,300 hectares and is dominated by western larch and Douglas-fir on the north facing slopes and ponderosa pine on south facing slopes, with a substantial intermixing of species. The eastern portion of the forest is represented by even-aged stands of lodgepole pine, with subalpine fir (*Abies lasiocarpa*) dominant in higher elevations.

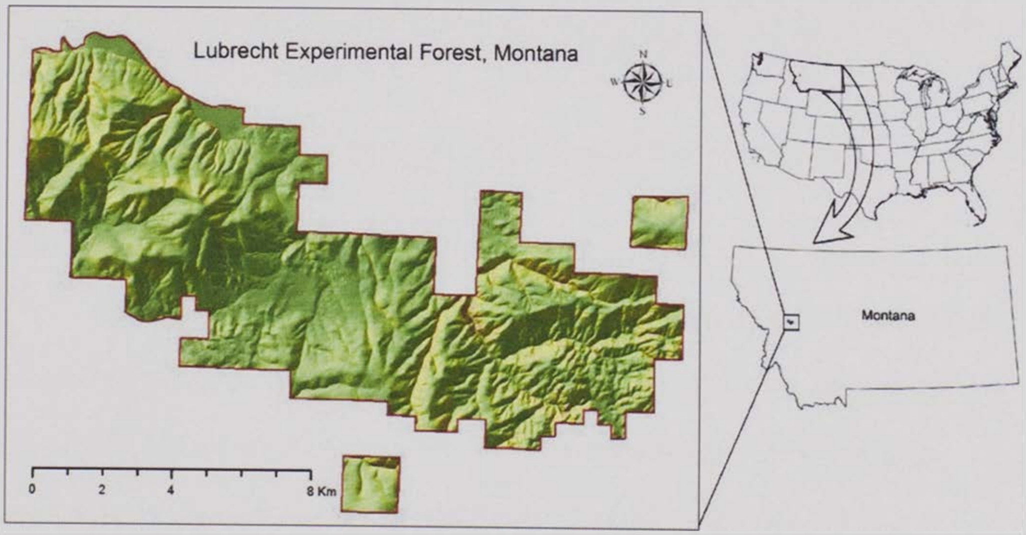
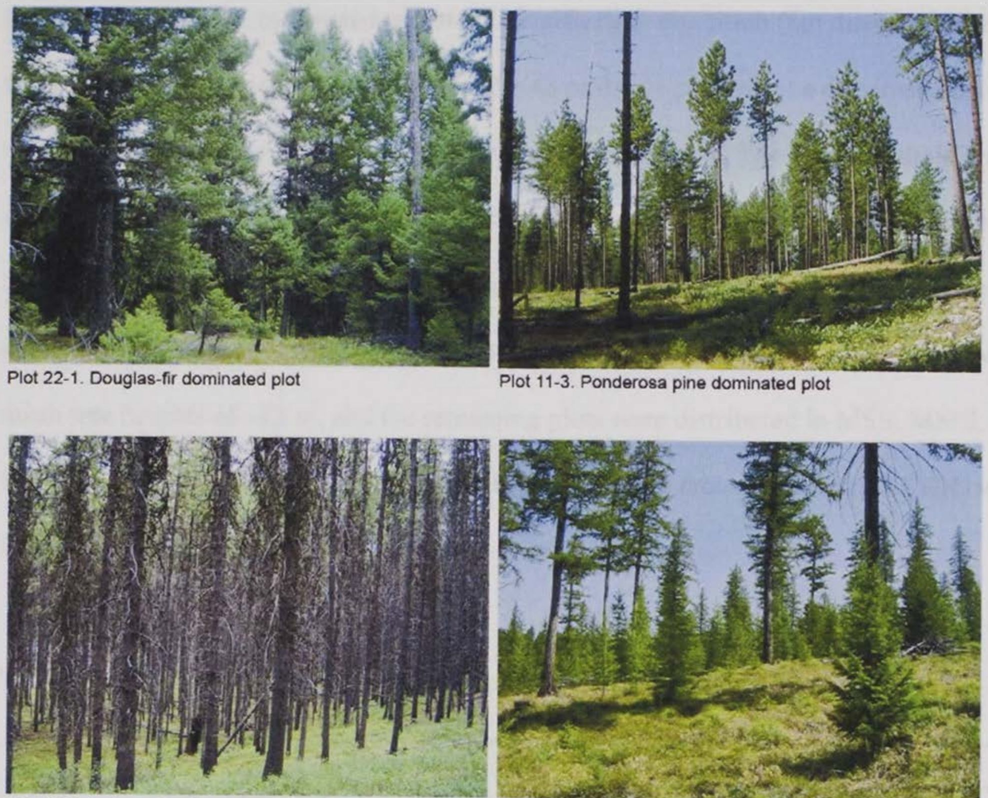


Figure 1. Lubrecht Experimental Forest, Montana.

In LEF, most ponderosa pine trees are represented by high open crowns with needle-like leaves in bundles of two or three that are green in color. The crown is irregular and flat although some trees are slightly conical near their tops (Figure 2). Douglas-fir typically has a pyramidal, symmetrical crown with needle-like leaves that are bluish-green in color. Mostly, the canopy is dense, depending on site conditions. Meanwhile, lodgepole pine is characterized by a long, slender trunk. The crown is thin on the top with needle-like leaves in bundles of two that are shorter than those of the ponderosa pine. In dense, even-aged stands, the canopy top is green while the middle and lower parts are dominated by dead branchwood. Western larch is the only deciduous conifer tree in the study area, characterized by an open, narrow conic crown with leaves clustered on branches that turn yellow in autumn and fall to the ground.



Plot 22-1. Douglas-fir dominated plot

Plot 11-3. Ponderosa pine dominated plot

Plot 12-8. Lodgepole pine dominated plot

Plot 15-8. Western larch dominated plot

Figure 2. Representative tree structures of the Lubrecht Experimental Forest, Montana.

Field data

A total of 61 rectangular plots (20 x 20 m^2) were established using a stratified random sample in two steps. First, lidar canopy height model (CHM) characteristics (canopy density and height variance) were used to generate five different canopy structures: (1) dense single strata stands (DSS), (2) dense multi strata (DMS), (3) moderate multi strata (MMS), (4) Moderate single strata (MSS), and (5) open (OPEN) (after Rowell et al. 2006). Second, a point generator placed on the study area was used to generate plots, which were subsequently stratified across the structure classes. It is worth acknowledging that the plots, on which 1555 trees were measured, were originally located to validate an automated tree stem identification algorithm (not directly related to this research). For the purposes of this study, 43 of the 61 plots meet a dominant species criteria, where dominant species is >70% on a tree count basis. The 70% threshold was selected to balance the need for within-plot species homogeneity against number of plots available for analysis. Four plots were located in DSS with an average of 60 trees/plot and mean tree heights of ~ 10 m, five plots were in DMS consisting ~50 trees/plot and mean tree heights of ~13 m, and the remaining plots were distributed in MSS, MMS, and OPEN as shown in Table 1. MMS has the highest average crownbase height (CBH) while DSS has a smaller mean height than other classes.

Table 1. Plot distribution and characteristics of canopy structures.

Structure Class	No. Plots	Canopy Cover	Type Canopy Cover	Height variance	Type Structure	Tree/plot	Height (m)			Crownbase (m)		
							Mean	Min.	Max.	Mean	Min.	Max.
Dense Single Strata	4	> 65%	Closed	≤ 20.25	Single Strata	59.75	9.89	7.96	11.83	5.38	3.22	6.23
Dense Multi strata	5	>65%	Closed	> 20.25	Multi strata	50.20	13.26	11.25	15.94	7.00	4.49	10.38
Moderate Single Strata	14	35-65%	Moderate Closed	≤ 20.25	Single Strata	39.14	13.12	10.11	17.11	6.71	2.69	11.16
Moderate Multi Strata	12	35-65%	Moderate Closed	> 20.25	Multi strata	22.17	19.11	12.36	25.88	9.83	5.94	14.33
Open	8	< 35%	Open	-	Open	7.38	16.69	7.76	22.72	8.56	1.80	13.19

Trees having diameter at breast height (DBH) > 7 cm were measured and canopy widths in two perpendicular directions (intersected approximately in the middle of a stem) were recorded. Trees with DBH < 7 cm were tallied by species. Tree heights and CBH were measured using a Forest Pro Laser with an effective accuracy of 0.3 m (Laser Technology, 2007) and tree positions were recorded by measuring range and distance from monumented plot corners, which were located using differentially corrected GPS measurements (average accuracy of ± 1.5 m). Additionally, each tree was classified using Kraft-class as dominant, codominant, intermediate, and suppressed (Oliver and Larson, 1996). At the dominant species level, 19 plots were represented by Douglas-fir with an average of 33 trees/plot followed by ponderosa pine (11) consisting of 21 trees/plot, western larch (7) and lodgepole pine (6). The range of tree heights within plots dominated by Douglas-fir was 8 - 26 m with a mean of ~14 m. Lodgepole pine plots contained mostly even-aged trees with a mean height of 13 m and CBH of 7 m as shown in Table 2. At the level of individual tree, 225 trees measured in the field were co-identified in the laser returns (process described below). Of these, Douglas-fir was the most abundant (84 trees), distributed within five different canopy structure classes with an average tree height of ~18 meters followed by western larch (63), lodgepole pine (44), and ponderosa pine (37). In general, western larch exhibited larger CBH than the other species.

Table 2. Description of trees within (a) dominant species and (b) individual tree species.

Species		No. Plots (trees)	Average trees per plot	Height (m)			DBH (Cm)			Crownbase (m)		
				Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.
Douglas-fir	(a)	19	32.47	14.25	7.96	25.88	22.47	11.42	40.68	7.49	3.22	14.33
	(b)	*84*	-	*17.89*	*5.10*	*30.20*	*29.58*	*7.00*	*64.60*	*9.31*	*0.30*	*18.20*
Ponderosa pine	(a)	11	21.18	16.32	11.43	23.59	25.58	16.74	39.21	8.26	4.15	13.19
	(b)	*41*	-	*18.28*	*4.20*	*28.80*	*30.21*	*9.20*	*77.47*	*9.18*	*0.40*	*15.20*
Lodgepole pine	(a)	6	42.83	13.69	10.11	17.11	17.86	12.69	25.20	6.82	2.69	11.16
	(b)	*37*	-	*15.55*	*7.00*	*20.40*	*21.14*	*8.00*	*41.40*	*8.53*	*1.70*	*14.80*
western larch	(a)	7	36.57	17.15	7.76	24.03	21.71	12.53	40.77	8.97	1.80	11.61
	(b)	*63*	-	*22.14*	*5.20*	*35.30*	*29.41*	*7.11*	*56.50*	*12.15*	*2.00*	*19.70*

Airborne laser scanning data acquisition

Airborne Laser Scanning data were obtained in July 27 and August 17, 2005 using Leica Geosystems Proprietary ALS50. The system was flown aboard a fix-wing aircraft at mean flying height of 1828 meters above mean terrain elevation. The data were scanned at nominal 1.5 meter post spacing across the track of the flight strips generating an irregular grid of data points on the ground. Information regarding post flight processing of the data and detailed parameter for this acquisition is shown in Table 1 of Chapter 2.

Methods

Delineation of individual trees

Following intensity normalization, individual trees were delineated using a stem identification algorithm (Rowell, et al., 2009) based on a combination of variable-window local maxima filtering (Popescu and Wynne, 2004) and neighborhood canopy height variance and return density (Rowell, et al. 2006). Returns were clipped from the CHM on plot boundaries. The LM method anticipates canopy width as a function of height (and stand structure) and searches a circle with dimensions of expected canopy width for points higher than the candidate point. If none are found, the candidate point is assumed to be a tree top. This evaluation is conducted on every point in the CHM to produce a stem map. It is worth noting that canopy overlap, as well as subtle variations in field and lidar stem geometries, complicates the assignment of returns to stems. For example, it is difficult to partition returns located within overlapping canopies of different tree species to a particular tree, potentially affecting intensity analysis

(Holmgren and Persson, 2004). Additionally, not all trees in the field can be linked to laser-detected stems due to stem detection inaccuracies and measurement offsets (Persson et al. 2002; Rowell et al. 2006). The stem detection algorithm used in this study produced a root mean square error (RMSE) of 17 stems per all plots (46.8%) for overstory and intermediate trees across all structure types (Rowell et al., 2009). Due to the issues described above, the following procedures were carried out to match individual field and laser stems and to assign returns to each of them.

Matching Field and Laser stems

Each detected tree was plotted on a map and tied to a corresponding field measured tree using a nearest neighbor approach. This was performed by calculating the closest field measured tree location corresponding to the detected stem (x,y-distance) with the maximum of 1.5 meter difference. Once a field and laser stem was matched, the estimated tree canopy diameter was used to assign laser returns to the individual tree. Returns from non-overlap canopy trees were automatically separated and assigned to corresponding trees, while returns from the overlapped canopies of two or more trees of the same species were tied to the tallest tree. Overlapping trees of different species were removed from consideration. Only trees with three or more laser returns at heights taller than 2.0 meters were used for intensity analysis to minimize erroneous inclusion of ground returns. A total of 225 out of 753 trees were selected with an average of 20 returns per tree. Only vegetation returns were used and classified into three types: "all," "first," and "single" returns. All returns were defined as the total of laser returns delineated within each tree crown. This included the first, second, and third returns.

Although few fourth returns were recorded, they were located at the lower canopy (< 2.0 meters in height) and automatically, were not included. The first returns were the first of multiple returns while the single returns were described as the only returns recorded for given pulses.

Statistical Analysis

Fourteen variables were calculated for plot-level dominant species and the same 14 variables were generated for individual trees (Table 3). They include: (1) percent of first and single returns (PCTF and PCTS), (2) means of canopy heights for all, first, and single returns (MCA, MCF, and MCS), (3) standard deviations of canopy heights for all return types (SDMCA, SDMCF, and SDMCS); (4) mean intensity of all, first and single returns (MIA, MIF, and MIS); and (5) standard deviation of mean intensity of all, first and single returns (SDMIA, SDMIF, and SDMIS).

A set of Exploratory Data Analyses (EDA) and three statistical analyses were performed. The EDA method uses the box-whisker plot to describe characteristics of intensity from dominant species class and individual tree species. The Pearson correlation analysis was performed following bivariate scatterplot inspections to evaluate relationships between the variables. This analysis was conducted to evaluate the degree to which pairs of variables are related. It also indicates how many variables are correlated each other and is used to help select the appropriate variables for the LDA, in conjunction with Principal Component Analysis (PCA). One-way ANOVA with Tukey's post hoc tests were used to evaluate differences in between species classes and to compute multiple comparisons simultaneously (14 pairwise comparisons of means), while maintaining significance at the 0.05 level. Fundamental knowledge about these tests is well documented in several publications (Quinn and Keough, 2002; Wheater and Cook, 2005; Warner, 2008). PCA was performed to explore variance, and to select multiple variables for analysis and the result was not presented.

Linear Discriminant Analysis (LDA) was performed to classify species at dominant and individual tree levels and to evaluate predictors (variables) of importance to class distinctions. LDA assumed that species class distribution was normally distributed and it was carried out with equal priori probabilities. Cross-validation (leave-one-out method) was used to assess the accuracy of the discriminant analysis due to a lack of independent data. Accuracy assessments were performed on a plot basis for dominant species and on a tree count basis for individual trees. Cohen's Kappa coefficients were also calculated to measure the agreement between the classifications. The Kappa class value suggested by Monserud and Leemans (1992) was used to rate the agreement as poor (0.40), fair (0.40-0.55), good (0.55-0.70), very good (0.7-0.85), and excellent (>0.85).

Table 3. Variables used in the analysis.

Acronym	Description
PCTF	Percent of first canopy returns
PCTS	Percent of single canopy returns
MIA	Mean intensity of all returns
SDMIA	Standard deviation of intensity of all returns
MCA	Mean canopy height of all returns
SDMCA	Standard deviation of canopy height of all returns
MIF	Mean intensity of first returns
SDMIF	Standard deviation of intensity of first returns
MCF	Mean canopy height of first returns
SCMCF	Standard deviation of canopy height of first returns
MIS	Mean intensity of single returns
SDMIS	Standard deviation of intensity of single returns
MCS	Mean canopy height of single returns
SDMCS	Standard deviation of canopy height of single returns

Results

Results are presented separately for dominant species and individual trees. Within each section, individual variables are summarized first, results of correlative analysis, ANOVA, and Tukey's are presented second, and the outcomes of Linear Discriminant Analysis and classification are presented last.

Dominant Species

Return types, heights and intensity

Table 4 shows the distribution of laser return-type proportions and the canopy heights of return types. Douglas-fir intercepts more single returns (44%) than the other three species while Ponderosa pine intercepts the fewest. Western larch exhibits a higher percentage of first returns (69%) than other species. In general, mean heights of lodgepole pine returns are lower than other species across return types while western larch has the highest mean heights.

Table 4. Laser return characteristics within plot-level dominant species class.

Species	Average laser returns per plot	Average percentage (%) return type			Heights of return type (meter)								
					All			First			Single		
		All	First	Single	Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.
Douglas-fir	331.68	100.00	44.03	43.76	12.55	2.00	29.77	12.33	2.02	28.82	14.28	2.00	29.77
Ponderosa pine	250.45	100.00	64.14	23.41	13.20	2.00	29.33	13.42	2.02	26.87	14.80	2.01	29.33
Lodgepole pine	389.50	100.00	55.54	32.56	10.07	2.01	21.00	10.04	2.01	20.52	11.66	2.01	21.00
western larch	430.71	100.00	69.81	24.89	16.43	2.03	29.51	17.16	2.09	27.21	19.18	2.03	29.51

The box-whisker plots illustrating the characteristics of intensity and standard deviation are presented in Figure 3. The plots indicate that the mean intensity of single returns is generally higher than other return types across all species. Douglas-fir (DF) is consistently represented by higher mean intensities for all return types than the other three species. The differences in mean intensities between lodgepole pine (LP), ponderosa pine (PP) and western larch (WL) are small as reflected by medians that fall between the 25th and 75th percentiles of the other species for all, first, and single returns. The box-whisker plots of the standard deviation of intensity show that Douglas-fir exhibits higher mean standard deviations for all and first returns while for the single return, the difference is within the range of other species. Lodgepole pine shows higher variability in the standard deviation of single returns than the other three species.

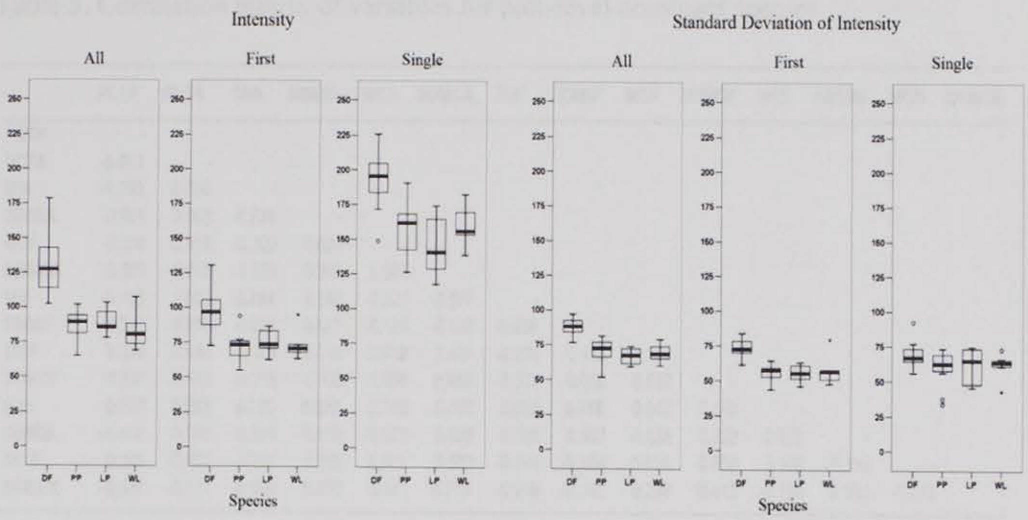


Figure 3. Box-whisker plots for plot-level dominant species for all, first, and single return types.

The Pearson correlations for 14 input variables are depicted in Table 5. The proportions of first (PCTF) and single returns (PCTS) are highly, negatively correlated (*r* < -0.70). However, both return type variables are weakly correlated with their corresponding mean intensities. All mean intensities are highly, positively correlated (r >0.70). The correlations between intensity and standard deviation vary, but in general all and first returns exhibit high correlation coefficients while single returns are weakly correlated (*r* <0.4). The three canopy heights and their standard deviations (SDMCA, SDMCF, and SDMCS) are weakly correlated with each other. Furthermore, most variables are not highly correlated, indicating independent behavior that may improve discrimination between species in Linear Discriminant Analysis (LDA). Consequently, all 14 variables were included in the LDA.

Table 5. Correlation matrix of variables for plot-level dominant species

	PCTF	PCTS	MIA	SDMIA	MCA	SDMCA	MIF	SDMIF	MCF	SDMCF	MIS	SDMIS	MCS	SDMCS
PCTF	-													
PCTS	***-0.911***	-												
MIA	***-0.750***	***0.786***	-											
SDMIA	-0.601	0.508	***0.734***	-										
MCA	0.355	-0.410	-0.303	0.037	-									
SDMCA	0.025	-0.160	-0.269	0.232	0.591	-								
MIF	-0.480	0.509	***0.892***	0.592	-0.221	-0.377	-							
SDMIF	-0.542	0.516	***0.855***	***0.837***	-0.138	-0.113	***0.858***	-						
MCF	0.376	-0.440	-0.366	0.010	***0.979***	0.684	-0.294	-0.193	-					
SDMCF	0.010	-0.132	-0.218	0.306	0.650	***0.829***	-0.333	-0.050	0.652	-				
MIS	-0.375	0.388	***0.725***	***0.839***	0.085	0.020	0.638	***0.718***	0.042	0.146	-			
SDMIS	-0.408	0.298	0.284	0.650	-0.077	0.339	0.198	0.492	-0.038	0.332	0.328	-		
MCS	0.396	-0.427	-0.336	0.020	***0.971***	0.506	-0.244	-0.164	***0.926***	0.650	0.102	-0.190	-	
SDMCS	-0.031	-0.111	-0.229	0.078	0.141	***0.775***	-0.315	-0.140	0.288	0.412	-0.189	0.364	-0.013	-

Species discrimination

Table 6 shows results of one-way ANOVA and Tukey's post hoc tests. The ANOVAs indicate significant differences between species for eight of 14 variables (*p-value* < 0.001) including percentages of first and single returns, mean intensities of all return types, and height of first returns. Tukey's post hoc tests reveal significant differences primarily between Douglas-fir and the other three species for seven variables. The differences in mean standard deviations of intensities are significant for all and first returns (SDMIA and SDMIF), but not for single returns (SDMIS) (*p-value* > 0.1). However, among three species (ponderosa pine, lodgepole pine, and western larch), the differences are not apparent, which also is visually indicated by the box-whisker plots (Figure 2). Although there is a significant difference in canopy heights (*p-value* <0.05) between the four species as represented by the ANOVA, Tukey's indicates that the difference occurs between Douglas-fir and western larch.

Table 6. The F statistic and P-values from one-way ANOVA for dominant species

Variable	ANOVA		Tukey's post hoc species difference (significant at the 0.05 level)			
	F	*P-value*	Douglas-fir (DF)	Ponderosa pine (PP)	Lodgepole pine (LP)	western larch (WL)
PCTF	19.44	<0.001	PP, LP, WL	DF	DF, WL	DF, LP
PCTS	11.39	<0.001	PP, WL	DF	-	DF
MIA	27.37	<0.001	PP, LP, WL	DF	DF	DF
SDMIA	23.65	<0.001	PP, LP, WL	DF	DF	DF
MCA	3.54	<0.023	-	-	WL	LP
SDMCA	2.31	<0.091	-	-	-	-
MIF	12.92	<0.001	PP, LP, WL	DF	DF	DF
SDMIF	21.88	<0.001	PP, LP, WL	DF	DF	DF
MCF	4.66	<0.007	WL	-	WL	DF, LP
SDMCF	1.74	<0.175	-	-	-	-
MIS	15.61	<0.001	PP, LP, WL	DF	DF	DF
SDMIS	2.10	<0.116	-	-	-	-
MCS	3.21	<0.033	-	-	WL	LP
SDMCS	0.96	<0.426	-	-	-	-

Classification

The range of classification accuracies using a single variable is 49-67%. A summary of classification accuracies produced by Linear Discriminant Analysis (LDA) is presented in Table 7. Only classifications resulting in accuracies greater than 70% are shown. Accuracies >70% are achieved using mean intensity of all returns (MIA) with either standard deviation intensity of all returns (SDMIA) or with mean canopy height of first returns (MCA). Accuracy consistently improves when percent return types (PCTF and PCTS) are included in the analysis. The maximum accuracy using variables selected from principal component analysis (PCA) is 93% with a cross-validated accuracy of 76%. However, a higher classification accuracy (95%) can be achieved when all variables are used in the analysis (cross-validated accuracy of 74%). In the latter case, more than 84% of the variance in all variables is explained.

Table 7. Variables used in the linear discriminant analysis and overall accuracy of classifications > 70%.

Variables	Accuracy (%)	
	Original grouped	Cross-validated
MIA, MCA	72.10	67.40
MIA, SDMCA	72.10	67.40
PCTF, PCTS, MIF, SDMIF	76.70	74.40
PCTF, PCTS, MIS, SDMIS	76.70	65.10
MIA, SDMIA, MIS, SDMIS, MCA, SDMCA	79.10	72.10
MIA, SDMIA, MCA, SDMCA, MIF, SDMIF, MCF	79.10	62.80
PCTF, PCTS, MIA, SDMIA	81.40	74.40
PCTF, PCTS, MIF, SDMIF, MIS, SDMIS	81.40	65.10
MIA, SDMIA, MCA, SDMCA, MIF, SDMIF	83.40	62.10
PCTF, PCTS, MIA, SDMIA, MIF, SDMIF, MIS	***83.70***	***79.10***
SDMIA, MCA, SDMCA, MIS, SDMIS, MCS, SDMCS	83.70	62.80
PCTF, PCTS, MIA, MIF, SDMIF	***86.00***	***79.10***
PCTF, PCTS, MIA, SDMIA, MIF, SDMIF, MIS, SDMIS	***86.00***	***74.40***
PCTF, PCTS, MIA, MIF, SDMIF, SDMCS	86.00	76.70
MIA, SDMIA, MIF, SDMIF, MIS, SDMIS, MCA, SDMCA, MCF, SDMCF, MCS, SDMCS	***88.40***	***69.80***
PCTF, PCTS, MIF, SDMIF, MIS, SDMIS, MCA, MCF, SDMCF, MCS, SDMCS	90.70	72.10
PCTF, PCTS, MIA, SDMIA, SDCMA, MIF, SDMIF, SCMCF, SDMCS	***93.00***	***76.70***
PCTF, PCTS, MIA, SDMIA, MIF, SDMIF, MIS, SDMIS, MCA, SDMCA, MCF, SDMCF, MCS, SDMCS	***95.30***	***74.40***

Table 8 shows the error matrix for the classification with the highest accuracy (95%). It is important to note that while the cross-validated classification is high the overall accuracy is reported using the original setting. The best producer's accuracies (the proportion of reference plots correctly classified) are for Douglas-fir and western larch (100%) followed by ponderosa pine (91%) and lodgepole pine (83%). The user's accuracy (the proportion of classified and reference plots that correctly correspond) is identical to the producer accuracy. Both evaluations (the producer and user) misclassify one plot of ponderosa pine and lodgepole pine to one another. The classification agreement is excellent as indicated by a Kappa value of greater than 0.90 (Monserud and Leemans, 1992).

Table 8. Error matrix resulting from discriminant analysis (the accuracy of 95%)

	Field reference				
	Douglas-fir	Ponderosa pine	Lodgepole pine	western larch	Row total
A. Classification					
Douglas-fir	**19**	0	0	0	19
Ponderosa pine	0	**10**	1	0	11
Lodgepole pine	0	1	**5**	0	6
western larch	0	0	0	**7**	7
Column total	19	11	6	7	**43**

B. Classification accuracy				
Producer's accuracy		User's accuracy		Overall accuracy = 95.35%
Douglas-fir	= 100.00%	Douglas-fir	= 100.00%	Kappa value = 0.93
Ponderosa pine	= 90.91%	Ponderosa pine	= 90.91%	
Lodgepole pine	= 83.33%	Lodgepole pine	= 83.33%	
western larch	= 100.00%	western larch	= 100.00%	

Individual trees

Return types, heights, and intensity

The distribution of return types and heights at the individual tree level is shown in Table 9. Ponderosa pine has the highest percentage of first returns (61%) while Douglas-fir has the lowest. More than 56% of returns are single returns for lodgepole pine while western larch has less than 24% single returns. Western larch is consistently represented by higher mean heights than the other three species (16 m, 17 m, and 19 m for all, first and single returns) while lodgepole pine is represented by the lowest.

Table 9. Laser return characteristics for individual trees.

Species	Average laser returns per tree	Average percentage (%) return type			Heights of return type (meter)								
					All			First			Single		
		All	First	Single	Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.
Douglas-fir	17.93	100.00	46.53	42.27	13.99	2.00	29.77	13.76	2.15	28.10	15.90	2.07	29.77
Ponderosa pine	18.17	100.00	60.68	27.06	14.00	2.00	30.24	14.19	2.04	28.98	16.05	2.08	30.24
Lodgepole pine	21.95	100.00	49.60	46.46	10.49	2.01	21.00	10.86	2.03	20.52	11.76	2.05	21.00
Western larch	20.54	100.00	56.98	23.10	15.97	2.04	35.12	16.65	2.09	31.81	19.04	2.04	35.12

Figure 4 shows box-whisker plots of intensities and standard deviations across species. Intensity is consistently higher for single returns than for all and first returns. Intensity is also highly variable across species classes, as indicated by minimum, maximum, and first and third percentile values. Douglas-fir has higher mean intensities than the other species for all, first and single returns. However, mean intensities of ponderosa pine, lodgepole pine and western larch overlap in range.

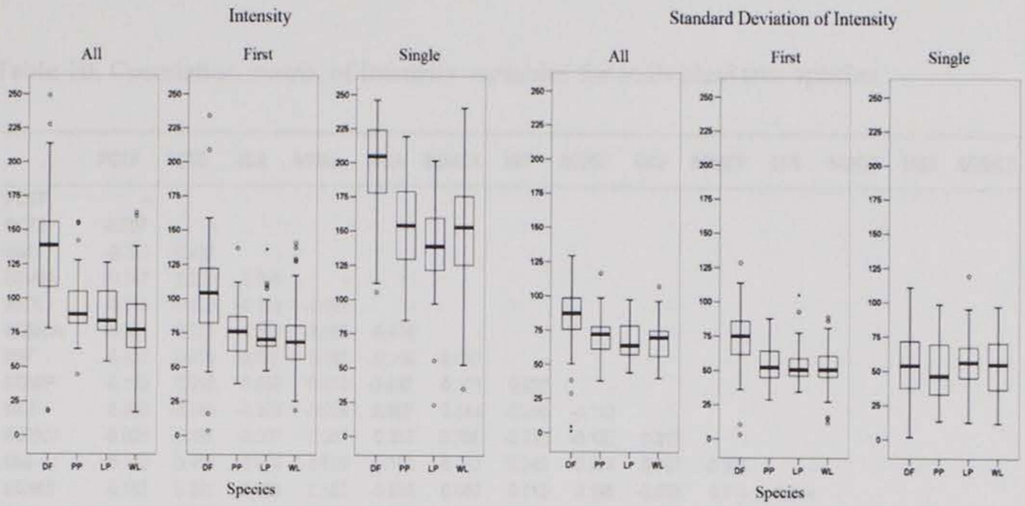


Figure 4. Box-Whisker plots of intensity and standard deviation for all, first, and single return types.

The bivariate correlations between variables are shown in Table 10. The percentage of first and single returns (PCTF and PCTS) are correlated ($r < -0.70$), but they do not show strong correlations with their corresponding mean intensity and height variables. Mean intensity is weakly correlated with its standard deviation, especially among single returns (MIS and SDMIS). The three mean canopy heights (MCA, MCF, and MCS) are highly, positively correlated to each other even though they are weakly associated with their standard deviations (SDMCA, SDMCF, and SDMCS). Additionally, intensity and canopy height variables are weakly negatively correlated. As with plot-level dominant species, most variables are not highly correlated, excepting the percentage of first and single returns (PCTF and PCTS). Each of the 14 variables is included in the modified LDA because they appear to contain different information.

Table 10. Correlation matrix of intensity variables for individual tree species.

	PCTF	PCTS	MIA	SDMIA	MCA	SDMCA	MIF	SDMIF	MCF	SDMCF	MIS	SDMIS	MCS	SDMCS
PCTF	-													
PCTS	***-0.707***	-												
MIA	-0.283	0.431	-											
SDMIA	-0.247	0.200	0.549	-										
MCA	0.114	-0.289	-0.119	-0.001	-									
SDMCA	-0.081	-0.221	-0.312	0.105	0.436	-								
MIF	-0.121	0.316	***0.751***	0.157	-0.160	-0.397	-							
SDMIF	-0.160	0.293	0.563	0.473	***-0.930***	-0.228	0.626	-						
MCF	0.098	-0.348	-0.203	-0.024	***0.947***	0.544	-0.210	-0.142	-					
SDMCF	-0.005	-0.167	-0.207	0.057	0.357	***0.704***	-0.320	-0.130	0.318	-				
MIS	-0.170	0.157	0.674	0.630	0.108	-0.012	0.345	0.344	0.067	-0.024	-			
SDMIS	-0.147	0.201	0.008	0.157	-0.086	0.053	0.113	0.195	-0.053	0.014	-0.244	-		
MCS	0.135	-0.323	-0.186	-0.010	***0.959***	0.495	-0.223	-0.148	***0.888***	0.404	0.101	-0.152	-	
SDMCS	-0.131	0.081	-0.194	0.019	-0.042	0.542	-0.169	-0.050	0.108	0.289	-0.152	0.403	-0.102	-

Species Discrimination

Differences between structure and intensity variables are presented in Table 11. With regard to ANOVA, most variables are significantly different between species (*p-value* < 0.001), excepting standard deviations of intensity and height for single returns (*p-value* > 0.01). The results of Tukey's post hoc tests are more illuminating in terms of highlighting where differences between each of the four species lie. For example, Douglas-fir is significantly different from the other three species in most of the intensity metrics (e.g., MIA, SDMIA, MIF, SDMIF, and MIS). Similar differences can be seen for the other tree species with different variables. For example, western larch is significantly different from the other species with respect to canopy height for first and single return types.

Table 11. The one-way ANOVA and Tukey's post hoc tests at the 0.05 significant level.

Variable	ANOVA		Tukey's post hoc species difference (significant at the 0.05 level)			
	F	*P-value*	Douglas-fir (DF)	Ponderosa pine (PP)	Lodgepole pine (LP)	western larch (WL)
PCTF	15.04	<0.001	PP, WL	DF, LP	PP, WL	DF, LP
PCTS	45.32	<0.001	PP, WL	DF, LP	PP, WL	DF, LP
MIA	45.56	<0.001	PP, LP, WL	DF	DF	DF
SDMIA	16.22	<0.001	PP, LP, WL	DF	DF	DF
MCA	16.29	<0.001	LP, WL	LP	DF, PP, WL	DF, LP
SDMCA	11.66	<0.001	WL	WL	WL	DF, PP, LP
MIF	22.26	<0.001	PP, LP, WL	DF	DF	DF
SDMIF	22.54	<0.001	PP, LP, WL	DF	DF	DF
MCF	19.06	<0.001	LP, WL	LP, WL	DF, PP, WL	DF, PP, LP
SDMCF	8.67	<0.001	LP, WL	WL	DF, WL	DF, PP, LP
MIS	35.53	<0.001	PP, LP, WL	DF	DF	DF
SDMIS	0.69	<0.561	-	-	-	-
MCS	15.75	<0.001	LP,WL	LP, WL	DF, PP, WL	DF, PP, LP
SDMCS	1.97	<0.120	-	-	-	-

Classification

The Linear Discriminant Analysis generates classification accuracies ranging from 37-52% with a single variable. Table 12 shows the results of classifications with accuracies exceeding 60%. Accuracies greater than 60% are achieved using intensity and height metrics for all returns (MIA and MCA). In contrast to the results of the dominant species level analysis, the percentage of return-type only slightly improves classification accuracy for individual trees. Accuracies of 68% (cross-validated 65%) are produced using six variables derived from all and first returns. With this model, slightly more than 82% of variance between species is explained. When all variables are used, the accuracy decreases slightly.

Table 12. Variable groups used in discriminant analysis and overall accuracy of classifications.

Variables	Accuracy (%)	
	Original grouped	Cross-validated
MIA, MCA	***60.40***	***59.10***
PCTF, PCTS, MIA, SDMIA	61.30	60.40
PCTF, PCTS, MIS, SDMIS	61.80	61.30
PCTF, PCTS, MIF, SDMIF, MIS, SDMIS	62.20	59.60
SDMIA, MCA, SDMCA, MIS, SDMIS, MCS, SDMCS	63.10	61.30
PCTF, PCTS, MIA, SDMIA, MIF, SDMIF, MIS, SDMIS	63.60	59.10
MIA, SDMIA, MCA, SDMCA	***65.30***	***64.90***
MIA, SDMIA, MIS, SDMIS, MCA, SDMCA	65.80	64.40
MIA, SDMIA, MCA, SDMCA, SDMIS	65.90	61.80
MIA, SDMIA, MCA, SDMCA, MIS	66.70	64.40
MIA, SDMIA, MCA, SDMCA, MIF, SDMIF, MCF	67.10	64.40
PCTF, PCTS, MIA, SDMIA, MCA, SDMIF, MCF, SDMCF, MIS, SDMIS, MCS	67.10	62.20
MIA, SDMIA, MIF, SDMIF, MIS, SDMIS, MCA, SDMCA, MCF, SDMCF, MCS, SDMCS	67.10	62.70
MIA, SDMIA, MCA, SDMCA, MIF, SDMIF	***67.60***	***65.30***
PCTF, PCTS, MIA, SDMIA, MCA, MCF, SDMCF, MIS, MCS	***68.00***	***63.60***

The error matrix for the highest accuracy species classification (68%) is shown in Table 13. Douglas-fir is best characterized while ponderosa pine is least well classified. All species are partially misclassified especially ponderosa pine and western larch, which often incorrectly classify to one another. Similar to plot-level dominant species, ponderosa pine is also incorrectly classified to lodgepole pine and vice versa. The Kappa value of 0.56 indicates that the classification agreement between laser variables and field data is good (Monserud and Leemans, 1992).

Table 13. Error matrix resulting from discriminant analysis for individual tree species.

	Field reference				
	Douglas-fir	Ponderosa pine	Lodgepole pine	western larch	Row total
A. Classification					
Douglas-fir	**65**	4	1	4	74
Ponderosa pine	8	**20**	9	15	52
Lodgepole pine	6	5	**26**	2	39
western larch	5	12	1	**42**	60
Column total	84	41	37	63	**225**

B. Classification accuracy				
Producer's accuracy		User's accuracy		Overall accuracy = 68.0%
Douglas-fir	= 87.80%	Douglas-fir	= 77.38%	Kappa value = 0.56
Ponderosa pine	= 38.46%	Ponderosa pine	= 48.78%	
Lodgepole pine	= 66.67%	Lodgepole pine	= 70.27%	
western larch	= 70.00%	western larch	= 66.67%	

Discussion

This research evaluates the usefulness of normalized ALS intensity and structure data for discriminating between four abundant tree species in a mixed conifer forest in western North America. The results indicate that discrimination is achievable using relatively sparse laser data typical of large-area, high-altitude acquisitions from fixed wing aircraft. Several points are drawn from the findings.

Dominant species

The variability in the proportions of laser returns types and mean canopy heights (Table 4) indicates that interaction of laser energy with canopy structural parameters is multi-dimensional, and that vertical and horizontal canopy variations affect laser interceptions. The significant difference in percentage of first and single returns between plot-level dominant species may be explained by differences in canopy structure within and between plots. For example, a low proportion of single returns in western larch is likely due to the open canopy form of this species. Small needle-like leaves clustered on sparse branches of western larch form an open canopy with a high crown base (Tables 1 and 2) increasing the likelihood that a single pulse will have multiple reflections. By contrast, Douglas-fir dominated plots occur within three strata (dense single strata, moderate multi strata and dense multi strata stands). The compact canopy of Douglas-fir in these strata is expected to intercept more single returns than first returns. For species that often have similar canopy structures, such as ponderosa pine and lodgepole pine, the observed differences in proportions of return types are likely influenced by tree density within the plots. Five of six plots on which lodgepole pine is dominant are located in the moderate single strata class with average tree densities higher than the ponderosa pine plots distributed across three canopy strata (moderate single strata, moderate multi strata, and OPEN).

With respect to intensity, the fact that mean intensity of single returns is higher than other return types across all species classes is intuitive, as pulses with multiple reflections show some partitioning of available energy between the reflections due to multiple scattering. The high intensity of Douglas-fir across return types is probably due to canopy structure differences described above as well as species specific differences. The dense, bluish-green canopy of Douglas-fir trees is likely to reflect higher intensity than other species. On the other hand, the low intensity returns of lodgepole pine may be attributed to the presence of abundant dead branchwood. Consequently, intensity may be comparatively reduced due to the lower reflectively of woody material versus leaves in the near infrared, or more scattering. Schreier et al. (1985) and Moffiet et al. (2005) found that intensity characteristics were affected by structural differences at both stands and individual tree levels. Schreier et al. (1985) noticed that stem density differences between Jack pine stands produced variability in return intensity. However, they also noted that observed variability was not consistent across their study area, finding that young Jack pine and adjacent Scots pine plantations exhibited similar intensity characteristics despite different tree densities. Moffiet et al. (2005) also noted that observed intensity differences between Poplar box and Cypress pine depended on stand structure in addition to individual tree attributes.

In Table 6, the most distinguishable species is Douglas-fir, which can be differentiated from the other three species using many different variables. In contrast, ponderosa pine can only be differentiated from Douglas-fir. The relatively low classification accuracies resulting from single variable analyses suggest that structural or intensity variables alone are not sufficient to distinguish one species from another at a plot-level. However, high accuracy is achieved using just two uncorrelated variables of intensity and height of all returns (Table 7). By comparison, any combination of intensity and height derived from first or single returns does not result in a high accuracy, even though the variables exhibit similar statistical characteristics of (1) being weakly correlated to each other as shown in Table 5, and (2) exhibiting a statistically significant difference (Table 6). These results indicate that all returns provide better discriminant power than first or single returns in describing variances between the four conifer species classes. Such indication is likely due to the statistical effect that all returns are represented by more samples than first and single returns. Additionally, improvements in classification accuracy obtained by including proportions of return types is expected, as Douglas-fir, ponderosa pine, lodgepole pine and western larch exhibit significant differences in those variables (Table 6). At the level of dominant species by plot, it appears that structural differences are as important as species-level intensity characteristics in the classification.

It is worth noting that many of the forest stands in the study area do not meet the dominant-species threshold of this research (>70% on a tree basis), raising questions about the ability to distinguish between stands with different proportions of species. Further, only two of 43 plots that met the dominant species criteria contained less than 80% dominancy on a tree basis (the rest were >80% single species). Of the two plots, one consisted of 79% ponderosa pine and 21% Douglas-fir while the other consisted of 73% Douglas-fir and 27% ponderosa pine. The former plot exhibited intensity values that centered on mean intensity of ponderosa pine (Figure 2). Similarly, the intensity values of the latter plot centered on the mean intensity of Douglas-fir. These two observations tentatively suggest that the 70% criterion chosen for this study reasonably defines dominant species. Simply, the 70% threshold is sufficient to represent species characteristics at a plot level.

Individual tree level

The results of the species classification for individual trees are slightly different than at plot-level dominant species. As shown in Table 9, the high proportion of single returns in lodgepole pine is likely due to the uniformly tight spacing of tree crowns characteristic of the species. Nearly 85% of lodgepole pine trees used in the analysis are from plots located in moderately dense single strata stands, which have a generally closed upper canopy. Conversely, more than 45% of Douglas-fir trees were selected from moderately dense multi strata stands, which are distinguished by higher height variance and a more open upper canopy, with more gaps for the laser to penetrate into the lower canopy. These trees share similarities with lodgepole pine in terms of proportions of return types. However, despite obvious stand structural differences, the intensity of returns is dissimilar between species. Returns from lodgepole pine are of relatively low intensity, suggesting that canopy structure is not the primary factor affecting intensity. Instead, low intensity may be due to the abundant dead branches attached to tree stems, a phenomenon also manifested at plot-level dominant species. On the other hand, Douglas-fir consistently produces the highest intensity returns, perhaps because of its dense, compact canopy and flat, green needles that are highly reflective in the near infrared. Clark et al. (2003) reported that Douglas-fir needles are more highly reflective at ~1000nm than lodgepole pine in the same forest region as this study. Anecdotally, then, intensity is not only affected by canopy closure, but also by species-specific characteristics. This finding corroborates previous findings of Moffiet et al. (2005) and Ørka et al. (2007) suggesting that canopy characteristics and foliage reflectivity both influence intensity, as do other stem elements such as branches and tree bark.

It is noteworthy that none of the 14 variables examined can be used alone to produce a classification of >70% at the individual tree level even though many of them are significantly different between species (Table 10). However, a classification accuracy of 52% using mean intensity of all returns suggests that aggregated returns of all types (first, single, multiple) has the best potential for classification. In addition, while the proportions of return types are important variables in classification at plot level dominant species, these variables only slightly improve discrimination at the individual tree level. This may be explained by a lack of information regarding the source of individual laser returns from a tree crown (e.g., whether from leaf, branch, stem or a combination), which must be highly variable. These results collectively suggest that structure and intensity characteristics derived from low density (~0.44/ m^2) laser data alone are not optimal for species discrimination at the individual tree level. Ørka et al. (2007), using high density (5 points/m^2) ALS data found that the intensity metric alone did not produce a high accuracy classification for spruce, birch, and aspen in the Scandinavian forests of Norway. Additionally, they found that a combination of intensities of first and second returns increased accuracy only slightly. In a similar study, Holmgren and Persson (2004) showed that a high accuracy could be achieved using a single variable (the return proportion or standard deviation of intensity) to classify Norway spruce, Scots pine, and deciduous trees. They also found that a combination of both variables produced a classification accuracy approaching 90%.

By comparison, the relatively low classification accuracies reported in our study for individual trees even when using many different combinations of laser metrics indicates high variability between trees of Douglas-fir, ponderosa pine, lodgepole pine and western larch, with considerable overlap (Figure 4). For example, the frequent misclassification of ponderosa pine and lodgepole pine occurs because both species share intensity and height characteristics. Holmgren and Persson (2004) also note that high variability between individual trees (both within similar species and between different species) contribute to lower classification accuracies.

While it should be noted that classifying the four conifer species used in our study is imperfect at the individual tree level, it is worth acknowledging that Douglas-fir can be readily differentiated from the other species, especially from lodgepole pine. Further, preliminary results using other classification schema suggest that the obstacles identified in this study are not insurmountable. However, there is much additional research needed before species can be mapped operationally in western North American mixed-conifer forests. For example, fuzzy classifications, based on biophysical setting, canopy structure, or species associations may be necessary to deal with the class ambiguities that are almost certain to occur. Additionally, *a priori* knowledge of species present on a site may provide rule-outs for ambiguous classifications. Combining multispectral data with ALS data may also improve species identification and would provide additional validation.

Conclusions

The research presented here confirms that low density laser scanner data (< 1 return/m^2) is useful for discriminating between Douglas-fir, ponderosa pine, lodgepole pine, and western larch in mixed conifer forest. The overall classification accuracies of 95% and 68% can be achieved at plot-level dominant species and for individual tree respectively. Douglas-fir is the most readily distinguishable species among the trees examined while ponderosa pine and lodgepole pine are often misclassified. In general, proportion of return type, height, and intensity metrics individually are not sufficient to produce high accuracy classifications. Instead, using combinations of at least two variables with all return types improves classifications at both levels. The importance of proportion of return type variables for improving classifications of the four tree species at the plot level does not translate into success in classifying individual trees. Such difference is probably due to the sensitivity of laser metrics to plot and tree level structural variations as well as to species-specific characteristics at both levels.

Acknowledgements

This study was funded by the National Center for Landscape Fire Analysis (NCLFA) in the College of Forestry and Conservation of the University of Montana and The USDA Forest Service McIntire-Stennis Program. I thank Eric Rowell, Crystal Stonesifer, Casey Teske, Erik Hakanson, Tim Wallace, Ann Hadlow, Josh Rodriquez, Martin Twer, and R.J. Hannah for field assistance. We also would like to thank Ron Roth (Leica Geosystems, Inc.) for his communications on technical aspects of ALS50 System. In addition, I thank Jim Riddering for his constructive comments and suggestions.

References:

Arno, S.F., Simmerman, D.G., Keane, R.E., 1985. Forest succession on four Habitat types in western Montana. General Technical Report INT-177. USDA Forest Service, Intermountain Forest and Range Experiment Station Ogden, Utah. 74pp.

Asner, G.P., Lobel, D.B., Law, B.E., 2002. View angles effects on canopy reflectance and spectral mixture analysis of coniferous forests using AVIRIS. International Journal of Remote Sensing 23 (11), 247-2262.

Brandtberg, T., 2007. Classifying individual tree species under leaf-off and leaf-on conditions using airborne lidar. ISPRS Journal of Photogrammetry and Remote Sensing 61 (5), 5-340.

Clark, R.N., Swayze, G. A., Wise, R., Livo, K. E., Hoefen, T.M., Kokaly, R. F., Sutley, S. J., 2003. USGS Digital Spectral Library splib05a. U.S. Geological Survey, Open File Report 03-395. http://pubs.usgs.gov/of/2003/ofr-03-395/ofr-03-395.html (accessed on January 12, 2008).

Cochrane, M.A., 2000. Using vegetation reflectance variability for species level classification of hyperspectral data. International Journal of Remote Sensing 21(10), 075-2087.

Donoghue, D.N.M., Watt, P.J., Cox, N.J., Wilson, J., 2007. Remote sensing of species mixtures in conifer plantations using LiDAR height and intensity data. Remote Sensing of Environment 110 (4), 509–522.

Gobakken, T., Naesset, E., 2004. Estimation of diameter and basal area distributions in coniferous forest by means of airborne laser scanner data. Scandinavian Journal of Forest Research 19 (6), 529-542.

Holmgren, J., Persson, A., 2004. Identifying species of individual trees using airborne laser scanning. Remote Sensing of Environment 90 (4), 415-423.

Korpela, I.S., 2008. Mapping of understory lichens with airborne discrete-return LiDAR data. Remote Sensing of Environment 112 (10), 3891–3897.

Laser Technology, 2007. Forest Pro Laser's guide. Centennial, Colorado.

Leica Geosystems, 2008. ALS50 Airborne Laser Scanner guide. Norcross, Georgia.

Maas, H.G., Bienert, A., Scheller, S., Keane, E., 2008. Automatic forest Inventory parameter determination from terrestrial laser scanner data. International Journal of Remote Sensing 29 (5), 1579-1593.

Means, J. E., Acker, S. A., Brandon, J. F., Renslow, M., Emerson, L., Hendrix, C. J., 2000. Predicting forest stand characteristics with airborne scanning lidar. Photogrammetric Engineering & Remote Sensing 66 (11), 1367-1371.

Moffiet, T., Mengersen K., Witte C., King R., Denham R., 2005. Airborne laser scanning: Exploratory data analysis indicates potential variables for classification of individual trees of forest stands according to species. ISPRS Journal of Photogrammetry and Remote Sensing 59 (2005), 289-309.

Monserud, R.A., Leemans, R., 1992. Comparing global vegetation maps with the Kappa statistic. Ecological Modeling 62 (4), 275-293

Nimlos, T. J., 1986. Soils of Lubrecht Experimental Forest. Miscellaneous Publication Number 44. Montana Forest and Conservation Experiment Station, Missoula, Montana, 41 p.

Oliver, C.D., Larson, B.C., 1996. Forest Stand Dynamics. Wiley, New York. 509 pp.

Ørka, H.O., Naesset, E., Bollandsas, O.M., 2007. Utilizing airborne laser intensity for tree species classification. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 3/W52), 300-304.

Persson, A., Holmgren, J., Söderman, U., 2002. Detecting and Measuring Individual Trees Using an Airborne Laser Scanner. Photogrammetric Engineering & Remote Sensing 68 (9), 925-932.

Popescu, S. C., Wynne, R. H., 2004. Seeing the trees in the forest: Using LIDAR and multispectral data fusion with local filtering and variable window size for estimating tree height. Photogrammetric Engineering & Remote Sensing 70 (5), 589−604.

Quinn, G.P., Keough, M.J., 2002. Experimental Design and Data Analysis for Biologists. Cambridge University Press. New York. 537pp.

Roth, R., 2007. Personal communication. Leica Geosystems, Inc.

Rowell, E., Seielstad, C., Vierling, L., Queen, L., Shepperd, W., 2006. Using laser Altimetry-base segmentation to refine automated tree identification in managed forests of the Black Hills, South Dakota. Photogrammetric Engineering & Remote Sensing 72 (12), 1379-1388.

Rowell, E., Seielstad, C., Queen, L., 2009. Implementation and Performance of a Local Maximum Tree Stem Identification Algorithm Using Vector-format Airborne Laser Scanner Data. Manuscript in preparation.

Schreier, H., Lougheed, J., Tucker, C., Leckie, D., 1985. Automated measurements of terrain reflection and height variations using an airborne infrared laser system. International Journal of Remote Sensing 6 (1), 101– 103.

Terrasolid, 2004. TerraScan user's guide, Helsinki.

Warner, R. M., 2008. Applied Statistics: From Bivariate Through Multivariate Techniques. Sage Publications, California. 1057pp.

Wheater, C. P, Cook, P., 2005. Using Statistics to Understand the Environment. Taylor & Francis, New York. 218pp.

Introduction

Tree species data are necessary in forest inventories to support various management activities including timber production, biomass estimation, wildlife habitat delineation, and site productivity prediction. Several studies have shown that airborne laser scanning (ALS) data can be used to identify tree species in different forest ecosystems (Holmgren and Persson, 2004; Ǿrka et al. 2007; Moffiet et al. 2005; Brandtberg et al. 2003; Brandtberg, 2007; Donoghue et al. 2007; Suratno et al., 2009). Holmgren and Persson (2004) and Ǿrka et al. (2007) used height and intensity with quadratic and linear discriminant analyses (QDA/LDA) to differentiate between deciduous and conifer trees in Scandinavian forests. Brandtberg et al. (2003) and Brandtberg, (2007) analyzed ALS-derived crown characteristics (leaf on/off) using ANOVA and LDA to show that oaks, red maples and yellow poplar in eastern hardwood forest of the United States exhibited distinct canopy structures. In Australian sub-tropical forests, Moffiet et al. (2005) demonstrated that white cypress pine and poplar box trees were separable by applying exploratory data analysis and LDA on vegetation permeability, mean height, and intensity of different return types. Most recently, a study conducted by Suratno et al. (2009) showed that combinations of means and standard deviations of normalized intensity and canopy height were useful for identifying four prominent conifer species (Douglas-fir, ponderosa pine, lodgepole pine, and western larch) in mixed forests of western Montana, U.S.A. The latter authors used ANOVA and LDA to show that Douglas-fir was distinguishable from the other species using intensity metrics alone, but discrimination between all four species required intensity and height combinations. Discriminatory power was considerably higher at plot-level (0.04ha) than for individual trees.

The consensus of previous work is that tree species identification is achievable in a range of environments using combinations of ALS-derived structure and intensity metrics with reported accuracies ranging from 67 to 95%. However, the aforementioned studies represent isolated examples at local scale with common, but constrained methodologies applied to specific species and ecosystem. In short, the classification and mapping of tree species at landscape scale using ALS-data has not yet been demonstrated in the literature. It is noteworthy that LDA with cross-validation has been utilized by many researchers to derive species classes but not applied to landscapes, and I anticipate considerable difficulty in the application of this technique to landscapes with interspersed species because many landscape features will not share characteristics with the classes used to create the discriminant functions. In addition, LDA is unable to select the features for classification when all class centroids are overlapping (Jieping et al., 2006). Consequently, models generated from discriminant functions may produce high error rates when attempting to allocate features to classes (Krzanowski, 1988).

An alternative method similar to LDA is maximum likelihood classification (MLC). While LDA is a statistical estimation requiring more steps prior to using it for classification with some consequence mentioned above, MLC is a classification technique widely used in the remote sensing community. MLC is based on the probability density function to calculate the likelihood that individual features belong to classes (Lillesand and Kiefer, 2000). One weakness of MLC is that every individual is assigned to a class regardless of distance from class mean. This becomes a processing time (efficiency) issue when many variables are involved (Jia and Richards, 1994). In order to overcome such problem, several studies propose reducing the number of variables using principle components analysis, supervised classification, or multi-stage classification (Jia and Richards, 1994; Guohui et al, 1999; Jia and Richards, 2003). A model based on LDA functions would reduce variables.

The research described in this chapter is aimed to demonstrate that the low density airborne laser scanning (ALS) data is useful to map species at landscape level. Objectives are: (1) to use the commonalities in metrics and methods of the previously cited work (Suratno et al., 2009) for mapping four conifer tree species, Douglas-fir, ponderosa pine, lodgepole pine and western larch in an 11,300 ha study area in western Montana, U.S.A at a 0.04 ha resolution, (2) to evaluate the consistency of produced maps using two different independent datasets, stand database (walkthrough) and inventory dataset (0.25 acre / 0.10 ha circular plots), and (3) to produce a classification model as simply as possible.

Approach

The efficacy of ALS data to classify landscape and tree species at different scales has been shown in chapters two and three respectively. The importance of structural and intensity variables were the main factors that distinguish ponderosa pine (*Pinus ponderosa*), Douglas-fir (*Pseudotsuga menziesii*), western larch (*Larix occidentalis*) and lodgepole pine (*Pinus contorta*). These variables are exploited in this project and applied with a slight modification of the methodology.

The difference between techniques used in this research and the previous analyses (Holmgren and Persson, 2004; Moffiet et al. 2005; Suratno et al., 2009) is that the separation of "mixed joint distribution" between classes is performed separately. For this procedure, a two-staged approach is used. First, Linear Discriminant Analysis (LDA)-based classification function coefficients are generated for twelve intensity and height variables and combined in a linear equation to produce a single layer representing the interaction of intensity and height as a function of species. This procedure is a logical extension of those analyses, in which LDA has been used to discriminate species. The percentages of return types (first and single returns) are not included. Instead, the percent canopy cover (PCC) is used to replace these two variables to reduce the number of variables producing a model as simple as possible.

Second, a standard supervised Maximum Likelihood Classification is applied using the modified LDA-species layer and PCC with equal prior probabilities of class samples. The rationale for including PCC is that in addition to reducing the variables, observed differences in intensity appear to be at least as much a function of canopy closure as species, and investigation of additional variables such as stem density confirms the relative importance of PCC. Moffiet et al. (2005) corroborate this observation, suggesting that canopy openness and spacing contribute greatly to variations in intensity of laser returns due to canopy fractional interception of laser energy within individual footprints.

Materials and Methods

Study Site

The Lubrecht Experimental Forest (LEF) located approximately 54 km northeast of Missoula, Montana covers 11,300 hectares with elevations from 1160 to 1930 m. Established in 1937 as a research, teaching and demonstration forest of Montana Forest and Conservation Experimental Station (MFCES), it is the center of forest educational development at the University of Montana. LEF is part of the Blackfoot River drainage lying next to 2,800 hectares of forests administered by the Montana Department of Natural Resources and Conservation. The major forest types is dominated by western larch and Douglas-fir on the north facing slopes and ponderosa pine on south facing slopes and both species largely mix in the lower well drained elevations. Lodgepole pine is found abundance as dense, even-aged stands all over the eastern part of the forest.

Field data collection

The field data used to validate the laser species-based classification was collected in summer 2006 and 2007 following FIREMON collection protocol (FIREMON, 2007). Sixty-one rectangular 0.04 hectare plots were collected from five canopy structures, including dense single strata, dense multi strata, moderate single and multi strata, and open strata. Only 43 of 61 plots contained dominant species >70% and were used for this study. Plot locations were determined using differentially corrected GPS measurements. The detail characteristics of the strata and plot distribution within these classes are described Chapter three.

Lidar data acquisition

Laser altimetry data were acquired for the entire Lubrecht Experimental Forest area encompassing more than 11,300 hectares. The campaign was designed to have 50% sidelap so that all flight strips covered the entire study area, including the Elk Creek watershed. In order to synchronize each point location recorded by both LIDAR and GPS and changes in aircraft positions (roll, pitch, and yaw) documented by IMU, post flight processing was performed by vendor (Horizon, Inc.) using proprietary software. Data were in LAS format and processed using Microstation Development Language (MDL) extension in TerraScan (Terrasolid, 2004). The detailed parameter for this acquisition is shown in Table 1 and the separation of bare earth and canopy returns is described in Section Materials and Methods of Chapter 2 and Appendix B.

Stand database

The Lubrecht Forest developed a stand database from combinations of serial intensive walk-through inventories and aerial photo interpretations in 1995 and updated it in 2000 (Waterman, 2000). Stand species is defined by a predominant (prime) species and one or more secondary (alpha) species, if the second species composes at least 10% of total canopy cover. Noting that there is ambiguity with the lack of quantitative species distribution data and the fact that some stands are not homogeneous, we assume that stand homogeneity is at least within the dominant species criteria (>70%) if the stand is represented by a prime species and no alpha species. Using this logic, the entire LEF was divided into 1005 stand polygons, comprised of 180 Douglas-fir stands (total areas of ~792 ha), 50 ponderosa pine stands (~257 ha), 102 lodgepole pine stands (~331 ha) and

12 western larch stands (~27 ha). The rest of the stands (661 polygons) were either "mixed" species or were dominated by subalpine fir (*Abies lasiocarpa*) or Engelman spruce (*Picea engelmanii*).

Inventory dataset

Field inventory data were collected in 2007 on a 200 acre grid (81 hectares) to complement the existing stand database (walk-through data) and to provide a comparison to imagery data and the LIDAR datasets used for this study. Data were collected on 0.25 acre/0.10 hectare circular permanent plots distributed across LEF by a University of Montana, College of Forestry and Conservation field crew. Large trees (DBH> 13 cm (5")) were measured for height, DBH, crownwidth, and crownbase height while saplings and seedlings were tallied within 1/50th and 1/100th acre subplots. Plots were located using recreational-grade GPS units and their centers were marked using orange-painted rebar stakes with metal cap tags containing plot number and cruise date. Adrados et al. (2002) demonstrated that mean location errors of non differential GPS units such as the one used in this data collection across a range of conditions and environments is 28 meters with standard deviation and maximum errors of 10 meters and 63 meters respectively. For the purpose of this study, 77 of 101 plots having trees > 13 cm DBH were classified into dominant species (>70%) on a tree basis. Of these plots, 27 were represented by Douglas-fir followed by ponderosa pine (11), lodgepole (7) and western larch (1), while 31 plots were mixed species (no dominant species).

Table 1. Plot distribution and characteristics of dominant species

Species	No. Plots	Height (m)			DBH (Cm)			Crownheight (m)		
		Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.
Douglas-fir	27	15.74	2.13	29.87	26.06	12.70	76.20	7.90	0.00	21.64
Ponderosa pine	11	18.40	3.05	32.31	31.13	12.70	76.20	8.90	0.00	19.20
Lodgepole pine	7	17.07	6.40	29.87	20.67	12.70	58.42	9.02	0.00	21.03
western larch	1	20.45	3.05	34.75	26.19	12.70	73.66	11.43	0.00	21.95

Methods

Creating modified LDA-based species and PCC data (stage 1)

All procedures were performed using regular grids with cell size corresponding to the size of the original field plots (20 x 20 m^2). Two grid datasets were created: (1) a modified Linear Discriminant Analysis (LDA)-species layer, and (2) Percent Canopy Cover (PCC). The modified LDA-species layer was created based on the tree species identification methodology described by Suratno et al. (2009), using the following steps: (1) twelve grid layers were created from the Canopy Height Model, representing means and standard deviations of intensity and canopy height for three return types (all returns, first returns of multiple returns, and single returns (only one return recorded for a given pulse)). Points lower than 2 meters in height were removed to avoid ambiguity between canopy and ground. The variables are listed in Table 3 Chapter three (both percentages of return types are excluded). (2) LDA with cross validation was performed using the previously cited variables on the 43 candidate field plots to classify the four tree species, resulting in 12 coefficients for each of four discriminant functions. For this study, the discriminant function representing Douglas-fir was used to create a single intensity-height interaction layer (called a modified LDA-species) by applying equation 1, below. The reason Douglas-fir coefficients were used is based on a previous result (Suratno et al., 2009) showing that Douglas-fir was the most easily distinguished species in the LDA classification.

$$\begin{aligned} \textit{Modified LDA-species} = {} & 0.97(MIA) + 1.85(SDMIA) - 29.64(MCA) - 32.79(SDMCA) + \\ & 0.45(MIF) + 0.06(SDMIF) + 18.38(MCF) + 27.76(SDMCF) - \\ & 0.37(MIS) - 0.47(SDMIS) + 12.55(MCS) + 12.45(SDMCS) - 138.32 \quad (1) \end{aligned}$$

The values generated from equation 1 (represented by pixels) were classified into four groups (1 group per species). The thresholds for each group are represented by the LDA mean value from each species discriminant function produced previously (Suratno, et al., 2009) plus or minus 10 points (DF: 136.69 $\pm$ 10, PP: 83.98 $\pm$ 10, LP: 74.82 $\pm$ 10, WL: 93.70 $\pm$ 10). The 10-point criterion represents approximately one standard deviation. Half of the criterion was used for overlapping thresholds between species. Values that did not meet threshold criteria were identified as a mixed species.

Additionally, a percent canopy cover (PCC) layer was generated from the lidar point cloud using the proportion of vegetation first returns to total bare earth and vegetation first returns. Both layers were co-registered and combined to generate a single multiband (LDA-PCC) dataset for the final classification.

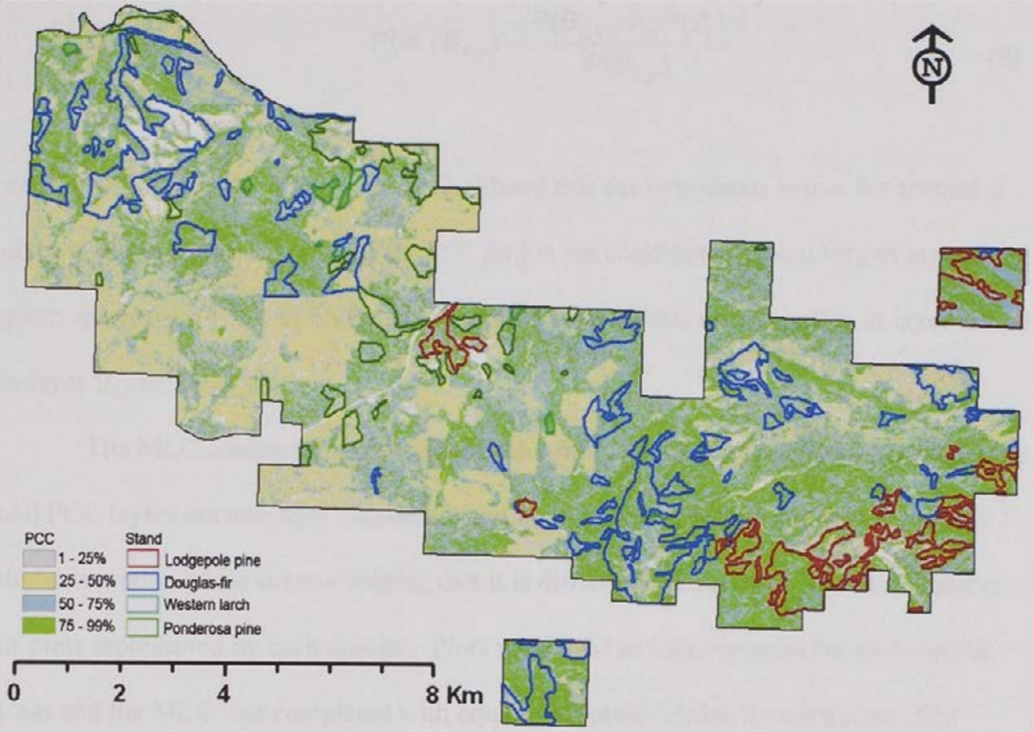


Figure 1. The distribution of percent canopy coverage (PCC) and 'dominant' species stands (Douglas-fir, ponderosa pine, lodgepole pine, and western larch) generated from the stand database. Everything not shown in colored polygons is mixed species.

Species classification and mapping (stage 2)

Classification and mapping were performed using a supervised Maximum Likelihood Classification (MLC) algorithm. The MLC is based on the probability that a pixel represented by species (A_i) with $i \in \{1, 2, \ldots, N \text{ species}\}$ located in x,y position of layer (B) is assigned into a particular class based on the pixel relative likelihood (probability) occurrence within the probability density function (P) of each species, which is defined by the Bayesian rule (Jensen and Nielsen, 2007) as:

$$P(A_i \mid B_{x,y}) = \frac{P(B_{x,y} \mid A_i)P(A_i)}{P(B_{x,y})} \quad (2)$$

where $P(A_i|B_{x,y})$ is the probability or likelihood that the hypothesis is true for species A given species distribution (layer) B, $P(B_{x,y}|A_i)$ is the conditional probability of layer B given species A, $P(A_i)$ and $(P(B_{x,y})$ are the prior probabilities of the species in layer A and in layer $B_{x,y}$ respectively.

The MLC assumes that species class distributions on both modified LDA-species and PCC layers are normally distributed and mutually independent and I tacitly accept this assumption while acknowledging that it is difficult to assess with the small number of plots represented by each species. Plots were used as training areas for each species class and the MLC was completed with equal prior probabilities for each class. The species classification was applied to the entire 11,300 ha study area.

Accuracy assessments were performed on a per pixel basis. Due to the ambiguity of the stand database quality and difference in resolutions between the map produced (0.04 hectare) and inventory plots (0.1 hectare), the following procedures were carried out for evaluating classification accuracy.

Accuracy assessment using stand database (stand polygons)

The proportion of pixels correctly classified (by species) to the total number of pixels in the analysis is calculated. The assessment is performed for the four species and for mixed species. The results are presented in two error matrices (1) classification accuracy for the four species independent of mixed species, and (2) classification accuracy including mixed species. Additionally, the accuracy assessment on mixed species is also conducted separately to evaluate the classification consistency given the fact that many stands are not pure/single species. This was performed by computing the proportion of pixels by species inside each stand. The species with the highest proportion was assumed as a "prime" and was used to match the category of prime species identified in the stand database for each stand. Only four prime species (Douglas-fir, ponderosa pien, lodgepole pine, and western larch) were used for matching while others comprising a total of 23 stands were excluded, including alpine fir (*Abies lasiocarpa*), cottonwood (*Populus deltoides*), aspen (*Populus tremuloides*), and Engelmann spruce (*Picea engelmannii*).

Accuracy assessment using a fixed permanent plot-based inventory dataset

Species classification accuracy was calculated using the 2007 fixed, permanent inventory plots to evaluate classification consistency using independent datasets. Two difficulties arise from us of such data. First, plot locations are uncertain due to coarse resolution GPS data (described above). Second, the resolution of the species map (0.04ha) does not match the resolution of the inventory plots (0.10ha). Therefore, the accuracy assessment was conducted based on proportions of classes/species (Woodcock and Strahler, 1987; Stehman and Czaplewski, 1998; Zhu et al., 2000), and the assessment was performed by calculating the proportion of pixels centered inside circles with radius of 63 meters. The radius of 63 meters was selected to accommodate the possible maximum observed location error generated by the recreational-grade GPS units (Adrados et al., 2002) used to establish each plot. The species with the highest proportion was assumed as the "dominant" species similar to the 70% dominant species criteria used previously. It is acknowledged that this definition of 'majority' is not the same as the dominant species criteria (70%) used in the previous analysis. For this reason, the assessment that results from this analysis is expected to provide only a broad overview of the consistency of species classification, and the ambiguity arising from the stated assumption is recognized.

All accuracy assessments are evaluated using Cohen's Kappa agreement. Coefficient values are calculated to indicate the agreement between the classifications. Monserud and Leemans (1992) suggested that categorized Kappa values result in the following five qualitative classes, (1) poor < 0.40, (2) fair = 0.40-0.55, (3) good = 0.55-0.70, (4) very good = 0.7-0.85, and (5) excellent >0.85.

Results

Results of the LDA and the MLC are depicted spatially in Figure 2, below. Species classification using modified LDA method produces an overall accuracy of 45% as shown in Table 2. The highest producer's accuracy (the total of correctly classified pixels from the class samples) is Douglas-fir (75%) followed by western larch, lodgepole pine, and ponderosa pine (39%, 21%, and 18% respectively). Meanwhile, the user's accuracies (the correctly assigned pixels to one specific class) of Douglas-fir, lodgepole pine, ponderosa pine, and western larch are 79%, 65%, and 3% respectively. The Kappa value is 0.21 indicating that the classification agreement between laser variables and field data is poor (Monserud and Leemans, 1992). The overall accuracy improves to 47% when mixed species is included in the analysis, but the Kappa value decreases to 0.07 (Table 3).

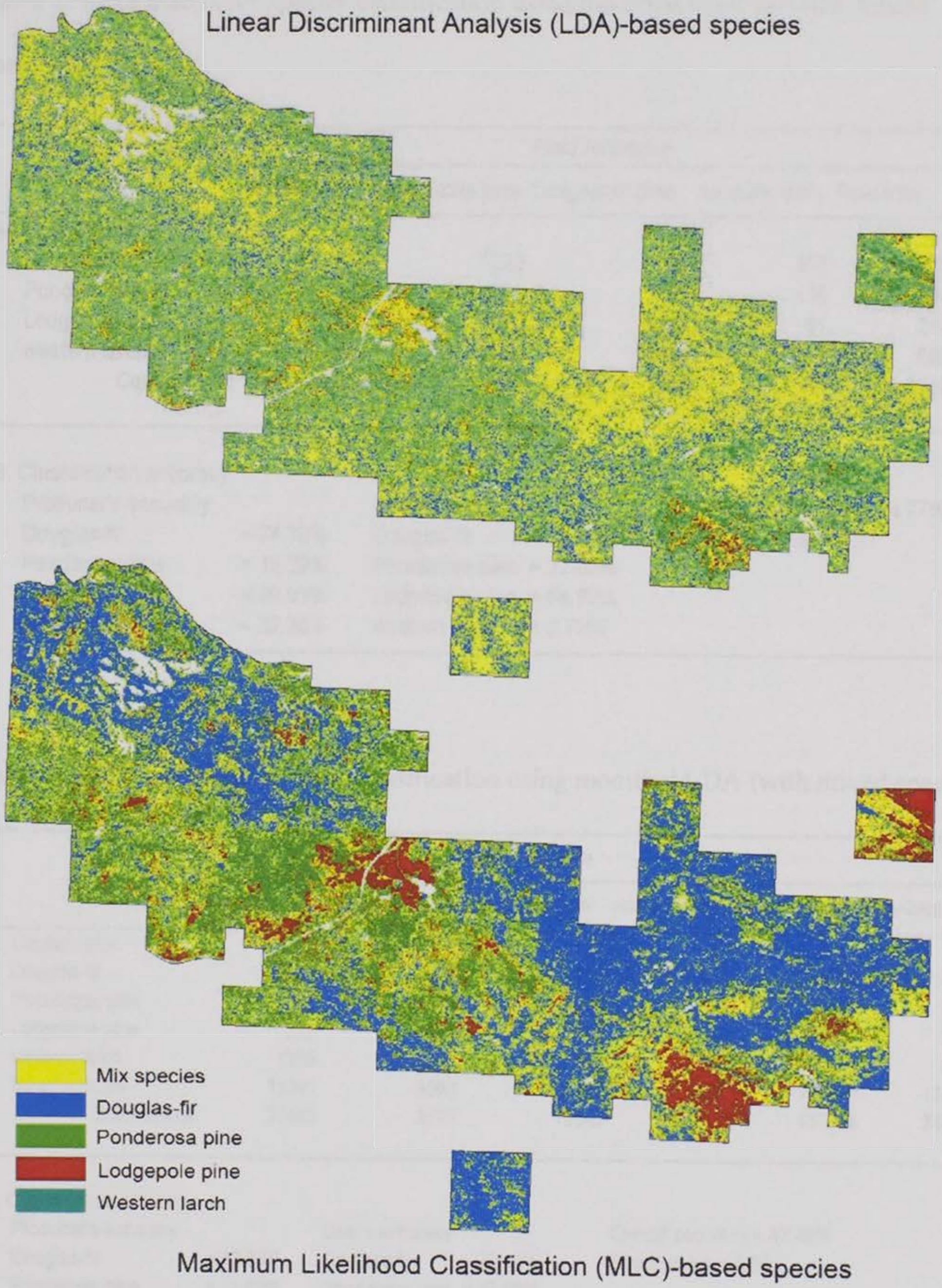


Figure 2. Species distribution based on modified Linear Discriminant Analysis (LDA) and Maximum Likelihood Classification (MLC).

Table 2. Error matrix for species classification using modified LDA (without mixed species).

	Field reference				
	Douglas-fir	Ponderosa pine	Lodgepole pine	western larch	Row total
A. Classification					
Douglas-fir	**8037**	1220	728	167	10152
Ponderosa pine	506	**931**	1899	110	3446
Lodgepole pine	222	575	**1607**	81	2485
western larch	1989	2364	3451	**218**	8022
Column total	10754	5090	7685	576	**24105**

B. Classification accuracy

Producer's accuracy		User's accuracy		
Douglas-fir	= 74.73%	Douglas-fir	= 79.17%	Overall accuracy = 44.77%
Ponderosa pine	= 18.29%	Ponderosa pine	= 27.02%	Kappa value = 0.21
Lodgepole pine	= 20.91%	Lodgepole pine	= 64.67%	
western larch	= 37.38%	western larch	= 2.72%	

Table 3. Error matrix for species classification using modified LDA (with mixed species).

	Field reference					
	Douglas-fir	Ponderosa pine	Lodgepole pine	western larch	Mixed species	Row total
A. Classification						
Douglas-fir	**8037**	1220	728	167	44294	10152
Ponderosa pine	506	**931**	1899	110	10021	3446
Lodgepole pine	222	575	**1607**	81	6163	2485
western larch	1989	2364	3451	**218**	29927	8022
Mixed species	16691	4083	5357	501	**107084**	133716
Column total	27445	9173	13042	1077	197489	**248226**

B. Classification accuracy

Producer's accuracy		User's accuracy		
Douglas-fir	= 42.56%	Douglas-fir	= 79.17%	Overall accuracy = 47.49%
Ponderosa pine	= 11.60%	Ponderosa pine	= 27.02%	Kappa value = 0.07
Lodgepole pine	= 12.85%	Lodgepole pine	= 64.67%	
western larch	= 20.24%	western larch	= 2.72%	
Mixed species	= 54.22%	Mixed species	= 80.87%	

Table 2. Error matrix for species classification using modified LDA (without mixed species).

	Field reference				
	Douglas-fir	Ponderosa pine	Lodgepole pine	western larch	Row total
A. Classification					
Douglas-fir	**8037**	1220	728	167	10152
Ponderosa pine	506	**931**	1899	110	3446
Lodgepole pine	222	575	**1607**	81	2485
western larch	1989	2364	3451	**218**	8022
Column total	10754	5090	7685	576	**24105**

B. Classification accuracy

Producer's accuracy		User's accuracy		
Douglas-fir	= 74.73%	Douglas-fir	= 79.17%	Overall accuracy = 44.77%
Ponderosa pine	= 18.29%	Ponderosa pine	= 27.02%	Kappa value = 0.21
Lodgepole pine	= 20.91%	Lodgepole pine	= 64.67%	
western larch	= 37.38%	western larch	= 2.72%	

Table 3. Error matrix for species classification using modified LDA (with mixed species).

	Field reference					
	Douglas-fir	Ponderosa pine	Lodgepole pine	western larch	Mixed species	Row total
A. Classification						
Douglas-fir	**8037**	1220	728	167	44294	10152
Ponderosa pine	506	**931**	1899	110	10021	3446
Lodgepole pine	222	575	**1607**	81	6163	2485
western larch	1989	2364	3451	**218**	29927	8022
Mixed species	16691	4083	5357	501	**107084**	133716
Column total	27445	9173	13042	1077	197489	**248226**

B. Classification accuracy

Producer's accuracy		User's accuracy		
Douglas-fir	= 42.56%	Douglas-fir	= 79.17%	Overall accuracy = 47.49%
Ponderosa pine	= 11.60%	Ponderosa pine	= 27.02%	Kappa value = 0.07
Lodgepole pine	= 12.85%	Lodgepole pine	= 64.67%	
western larch	= 20.24%	western larch	= 2.72%	
Mixed species	= 54.22%	Mixed species	= 80.87%	

Table 4 shows the classification using MLC method. The method generates overall accuracy of 76% with the highest producer's accuracy for Douglas-fir (88%) followed by lodgepole pine and ponderosa pine (68% and 56% respectively). However, class western larch produces an accuracy of 0%. Meanwhile, the user's accuracies of Douglas-fir, lodgepole pine, and ponderosa pine are 84%, 76%, and 50%. Identical to the producer's accuracy, class western larch generates an accuracy of 0%. In contrast to the Kappa value of the modified LDA classification, the value resulting from the MLC is 0.59 suggesting that the agreement is good. However, adding mixed species results in lower overall accuracy (36%) and a lower Kappa value of 0.11 (poor) (Table 5).

Table 4. Error matrix for species classification using MLC (without mixed species).

	Field reference				
	Douglas-fir	Ponderosa pine	Lodgepole pine	western larch	Row total
A. Classification					
Douglas-fir	**17072**	1859	1000	298	20229
Ponderosa pine	1666	**3462**	1639	99	6866
Lodgepole pine	578	916	**5594**	289	7377
western larch	1	1	1	**0**	3
Column total	19317	6238	8234	686	34475

B. Classification accuracy				
Producer's accuracy		User's accuracy		Overall accuracy = 75.79%
Douglas-fir	= 88.38%	Douglas-fir	= 84.39%	Kappa value = 0.59
Ponderosa pine	= 55.50%	Ponderosa pine	= 50.42%	
Lodgepole pine	= 67.94%	Lodgepole pine	= 75.83%	
western larch	= 0.00%	western larch	= 0.00%	

Table 5. Error matrix for species classification using MLC (with mixed species).

	Field reference					
	Douglas-fir	Ponderosa pine	Lodgepole pine	western larch	Mixed species	Row total
A. Classification						
Douglas-fir	**17072**	1859	1000	298	90458	20229
Ponderosa pine	1666	**3462**	1639	99	33545	6866
Lodgepole pine	578	916	**5594**	289	11288	7377
western larch	1	1	1	**0**	9	3
Mixed species	8128	2937	4821	391	**62174**	78451
Column total	27445	9175	13055	1077	197474	**248226**

B. Classification accuracy				
Producer's accuracy		User's accuracy		Overall accuracy = 35.57%
Douglas-fir	= 88.38%	Douglas-fir	= 84.39%	Kappa value = 0.11
Ponderosa pine	= 55.5%	Ponderosa pine	= 50.42%	
Lodgepole pine	= 67.94%	Lodgepole pine	= 75.83%	
western larch	= 0.00%	western larch	= 0.00%	
Mixed species	= 31.48%	Mixed species	= 79.25%	

Recall that for all stands that did not have a dominant species (as defined above), the accuracy assessment was performed on a per stand basis in which the majority species (highest proportion) as indicated by the lidar data was compared to the prime species of the walkthrough inventory. These results are shown in Table 6. The overall accuracy is 66% and the highest producer's accuracy is for Douglas-fir (88%) followed by ponderosa pine and lodgepole pine (36% and 30% respectively). Meanwhile, western larch produces an accuracy of 0%. It is important to note that the accuracy assessment presented here is limited to stands with a prime species and a different alpha species. Stands with the same prime and alpha species were classified as dominant in the analysis presented above and removed from further consideration. Additionally, the classification accuracy is performed based on the result of MLC alone.

Table 6. Error matrix for mix species classification using MLC (based on stand polygon).

	Field reference				
	Douglas-fir	Ponderosa pine	Lodgepole pine	western larch	Row total
A. Classification					
Douglas-fir	**366**	71	28	45	510
Ponderosa pine	40	**42**	5	12	99
Lodgepole pine	9	5	**14**	1	29
western larch	0	0	0	**0**	0
Column total	415	118	47	58	638

B. Classification accuracy				
Producer's accuracy		User's accuracy		Overall accuracy = 66.14%
Douglas-fir	= 88.19%	Douglas-fir	= 71.76%	Kappa value = 0.24
Ponderosa pine	= 35.59%	Ponderosa pine	= 42.42%	
Lodgepole pine	= 29.79%	Lodgepole pine	= 48.28%	
western larch	= 0.00%	western larch	= 0.00%	

Tables 7 and 8 depict the classifications using MLC method evaluated by the fixed-plot inventory dataset. The overall classification accuracy is 85% when mixed species are not included. It decreases to 60% when mixed species are added to the calculation. Both assessments indicate that Douglas-fir exhibits the highest producer and user accuracies. On the other hand, ponderosa pine and lodgepole pine produce accuracies of 0%. Meanwhile, more than half of mixed species are correctly classified (55%). The Kappa value of 0.61 indicates that the classification agreement is considered good when mixed species is excluded. As mixed species are added (Table 8), the Kappa value declines to 0.44 suggesting that the classification agreement between laser variables and field inventory data is fair (Monserud and Leemans, 1992).

Table 7. Error matrix for species classification using MLC (based on inventory dataset without mixed specie).

	Field reference				
	Douglas-fir	Ponderosa pine	Lodgepole pine	western larch	Row total
A. Classification					
Douglas-fir	**25**	2	0	0	27
Ponderosa pine	0	**4**	2	0	6
Lodgepole pine	0	1	**0**	0	1
western larch	0	0	0	**0**	0
Column total	25	7	2	0	34

B. Classification accuracy

Producer's accuracy		User's accuracy		
Douglas-fir	= 100.0%	Douglas-fir	= 92.59%	Overall accuracy = 85.29%
Ponderosa pine	= 57.14%	Ponderosa pine	= 66.67%	Kappa value = 0.61
Lodgepole pine	= 0.00%	Lodgepole pine	= 0.00%	
western larch	= 0.00%	western larch	= 0.00%	

Table 8. Error matrix for species classification using MLC (based on inventory dataset with mixed species included).

	Field reference					
	Douglas-fir	Ponderosa pine	Lodgepole pine	western larch	Mixed species	Row total
A. Classification						
Douglas-fir	**25**	2	0	0	10	37
Ponderosa pine	0	**4**	2	0	2	8
Lodgepole pine	0	1	**0**	0	2	3
western larch	0	0	0	**0**	0	0
Mixed species	2	4	5	1	**17**	29
Column total	27	11	7	1	31	77

B. Classification accuracy

Producer's accuracy		User's accuracy		
Douglas-fir	= 92.59%	Douglas-fir	= 92.59%	Overall accuracy = 59.74%
Ponderosa pine	= 36.36%	Ponderosa pine	= 66.67%	Kappa value = 0.44
Lodgepole pine	= 0.00%	Lodgepole pine	= 0.00%	
western larch	= 0.00%	western larch	= 0.00%	
Mixed species	= 54.84%	Mixed species	= 58.62%	

Discussion

The results above suggest that it may be possible to use relatively low density airborne laser scanning data to classify and map trees in mixed coniferous forests like those found at LEF. Application of the modified LDA function results in a grainy distribution of species that falsely emphasizes western larch and mixed classes (Figure 2). MLC resolves much of this ambiguity while improving the discrimination of lodgepole pine and ponderosa pine. Note that western larch is not apparent in the MLC classification and instead, grouped with ponderosa pine and Douglas-fir. The fractions of mixed species also decrease from modified LDA to MLC.

However, it is noteworthy that classification accuracy of MLC is low when mixed species are considered and there is still uncertainty about how well the approach captures stand heterogeneity. Yet, the mixed validation fails because the classification produces few mixed pixels relative to species-specific pixels. Therefore, during the validation, we are comparing species pixels (classified) to mixed pixels (truth) and can only conclude that we misclassify most of the time. Meanwhile, validation at stand-level instead of pixel level used in this study is also difficult as to whether the highest proportion of species is representative throughout each stand. This occurs particularly in stands having slight differences in pixel proportions between dominant mixed species (prime). For this reason, the validation presented in Table 6 is intended to provide a sense of classification consistency, which may provide useful knowledge for further studies.

The accuracy assessment (Table 6) shows that Douglas-fir is consistently classified with higher accuracy on both the producer and user sides. This result suggests that Douglas-fir can be located based on the ALS data in mixed settings. The genus *Pinus* is also distinctive, but differences are not readily apparent at the species level (e.g., ponderosa pine and lodgepole pine often are confused for one another). The higher classification accuracy for Douglas-fir is probably due to the larger number of samples distributed across diverse stands, and the low accuracy for western larch is conversely attributable sample-size constrained misrepresentation of species variability. Because the MLC uses probability to apportion pixels into classes, a representative probably density function for each attribute (per species) is essential. Additionally, western larch almost always occur secondary to ponderosa pine and Douglas-fir (Arno, et al. 1985). The high classification accuracy of lodgepole pine is likely due to the fact that laser return intensity is generally lower than for other species (Suratno et al., 2009) and the species usually occurs in even-aged stands.

Similarly, the higher percentage of Douglas-fir dominated stands (within mixed species stands) may be attributed to the general condition of species composition in Lubrecht Experimental Forest (LEF). Most mixed stands are largely composed by Douglas-fir as a seral species, especially in lower and middle elevation in LEF. On the other hand, western larch dominated stands are under-represented, resulting in an inadequate probability function in the MLC.

The high classification accuracy (85%) produced using the fixed plot inventory dataset for comparison suggests that the MLC methodology used to map species has the potential for larger application. Additionally, as was previously noted, percent canopy cover strongly affects return intensity due to fractional interception of energy by scattered foliage within the laser footprint. Consequently, the addition of PCC to the classification is logical and improves the result. We note that PCC was included as a variable in the

LDA with little accuracy improvement to the modified LDA classification. We speculate that the reason PCC improves the MLC classification but not the modified LDA classification is because the interaction between intensity and PCC is non-linear. Additionally, the MLC is not constrained in the same way and instead relies on probabilities.

Conclusions

The research presented here represents a first attempt at classification with promising results and suggests that ALS-data can be used to classify and map at least three western North American conifer species. The Maximum Likelihood Classification (without the mixed species) produced an overall accuracy of 75% with the best results for Douglas-fir followed by lodgepole pine and ponderosa pine, while western larch was difficult to identify. The ambiguity of species distributions in the stand database used for ground truth in mixed forest prevented robust validation for a large fraction of the landscape. However, the classification resulted in the correct species appearing in each stand polygon and a stand based validation using the prime/alpha notation indicated a qualitative agreement between the lidar classification and the walkthrough inventory data. It also produces a high accuracy as evaluated using independent data (fixed plot inventory dataset). The inclusion of prior probability weights informed by biophysical attributes such as slope, aspect, and elevation may clear up some of the uncertainty between species that exhibit strong site preferences. Additionally, it may be possible to improve the classification by increasing the number of training samples to better represent variability. We note that the samples were selected based on structure type rather than species. Lastly, we highlight the simplicity and repeatability of methods presented, but acknowledge considerable uncertainty regarding classification performance in mixed species stands.

Acknowledgements

This study was funded by a McIntire-Stennis Forestry Research Grant and by the National Center for Landscape Fire Analysis (NCLFA) in the College of Forestry and Conservation at the University of Montana. I thank Eric Rowell, Crystal Stonesifer, Casey Teske, Erik Hakanson, Tim Wallace, Ann Hadlow, Josh Rodriquez, Martin Twer, and R.J. Hannah for field assistance. We also would like to thank Ron Roth (Leica Geosystems, Inc.) for his communications on technical aspects of the ALS50 System.

References:

Adrados, C., Girard, I., Gendner, J.P., Janeau, G., 2002. Global Positioning System Location accuracy improvement due to Selective Availability removal. Comptes Rendul Biologies 325 (2002), 165-170.

Arno, S.F., Simmerman, D.G., Keane, R.E., 1985. Forest succession on four habitat types in Western Montana. General Technical Report INT-177. USDA Forest Service, Intermountain Forest and Range Experiment Station Ogden, Utah.

Brandtberg, T., Warner, T.A., Landenberger, R.E., McGraw, J.B., 2003. Detection and analysis of individual leaf-off tree crowns in small footprint, high sampling density lidar data from the eastern deciduous forest in North America. Remote Sensing of Environment, 85, 290–303.

Brandtberg, T., 2007. Classifying individual tree species under leaf-off and leaf-on conditions Using airborne lidar. ISPRS Journal of Photogrammetry and Remote Sensing, 61(5): 325-340.

Chen, J.M., Pavlic, G., Brown, L., Cihlar, J., Leblanc, S.G., White, H.P., Hall, R.J., Peddle, D.R., King, D.J., Trofymow, J.A., Swift, E., Van der Sanden, J., Pellikka, P.K.E., 2002. Derivation and validation of Canada-wide coarse-resolution leaf area index maps using high-resolution satellite imagery and ground measurements. Remote Sensing of Environment, 80, 165–184.

Donoghue, D.N.M., Watt, P.J., Cox, N.J., Wilson, J., 2007. Remote sensing of species Mixtures in conifer plantations using LiDAR height and intensity data. Remote Sensing of Environment,110, 509-522.

Fazakas, Z., Nilson, M., Olsson, H., 1999. Regional forest biomass and wood volume using satellite data and ancillary data. Agricultural and Forest Meterology, 98-99: 417-425.

FIREMON, 2007. Fire Effects Monitoring and Inventory System. http://frames.nbii.gov/ffi (accessed on June 16, 2007).

Franco-Lopez, H., Ek, A. R., Bauer, M. E., 2001. Estimation and mapping of forest stand density, volume, and cover type using the k-nearest neighbors method. Remote Sensing of Environment,77, 251-274.

Guohui, H., Desai, M.D., Zhang, X., 1999. Minimum component eigen-vector based classification technique with application to TM images. IEEE International Conference on Acoustics, Speech, and Signal Processing, 15-19 March 1999. Phoenix, Arizona.

Holmgren, J., Persson, A., 2004. Identifying species of individual trees using airborne laser scanning. Remote Sensing of Environment 90 (4), 415-423.

Jensen, F.V., Nielsen, T.D., 2007. Bayesian Networks and Decision Graphs. Second Edition. Springer. New York. 447 p.

Jia X., Richards, J.A., 1994. Efficient maximum likelihood classification for imagingspectrometer datasets. IEEE Transactions on Geoscience and Remote Sensing, 32(2): 274-281.

Jia X., Richards, J.A., 2003. Efficient transmission and classification of hyperspectral image data. IEEE Transactions on Geoscience and Remote Sensing, 41(5): 1129-1131.

Jieping Y., Xiong, T., Janardan, R., 2006. CPM: A Covariance-preserving Projection Method. Proceedings: Siang Conference on Data Mining 2006. Bethesda, Maryland, April 20-22, 2006.

Korpela, I.S., 2008. Mapping of understory lichens with airborne discrete-return LiDAR data. Remote Sensing of Environment 112 (10), 3891–3897.

Krzanowski, W.J., 1988. Principles of Multivariate Analysis: A User's Perspective. Clarendon Press. Oxford. 563 pp.

Lillesand, T.M., Kiefer, R.W., 2000. Remote Sensing and Image Intepretation 4th Edition, John Wiley & Sons, New York 724 p.

Lunetta, R.S., Lyon, J.G., 2004. Remote sensing and GIS accuracy assessment. CRC Press, New York. 304 pp.

Moffiet, T., Mengersen K., Witte C., King R., Denham R., 2005. Airborne laser scanning: Exploratory data analysis indicates potential variables for classification of individual trees of forest stands according to species. ISPRS Journal of Photogrammetry and Remote Sensing 59 (2005), 289-309.

Monserud, R.A., Leemans, R., 1992. Comparing global vegetation maps with the Kappa statistic. Ecological Modeling 62 (4), 275-293

Ørka, H.O., Naesset, E., Bollandsas, O.M., 2007. Utilizing airborne laser intensity for tree species classification. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 3/W52), 300-304.

Persson, A., Holmgren, J., Söderman, U., 2002. Detecting and Measuring Individual Trees Using an Airborne Laser Scanner. Photogrammetric Engineering & Remote Sensing 68 (9), 925-932.

Stehman, S. V., Czaplewski, R. C., 1998. Design and analysis for thematic map accuracy assessment: fundamental principles. Remote Sensing of the Environment. 64: 331–344.

Suratno, A., C. Seielstad, L. Queen, in press. Tree Species Identification in Mixed-Coniferous Forest Using Airborne Laser Scanning. ISPRS Journal of Photogrammetry and Remote Sensing (2009).

Terrasolid, 2004. TerraScan user's guide, Helsinki.

Wang, G., Gertner, G., Xiao, X., Wente, S., Anderson, A.B., 2001. Appropriate Plot Size and Spatial Resolution for Mapping Multiple Vegetation Types Photogrammetric Engineering & Remote Sensing, 67(5): 575-584.

Waterman, J.R., 2000. Building a Lubrecht forest inventory internet mapping service and volume comparisons. The Univeristy of Montana. Missoula, MT.

Woodcock, C. E., Strahler, A. H., 1987. The factor of scale in remote sensing. Remote Sensing of the Environment. 21: 311–332.

Zhu, Z., Yang, L., Stehman, S. V., Czaplewski, R. C., 2000. Accuracy assessment for the U.S. Geological Survey Regional Land-Cover Mapping Program: New York and New Jersey Region. Photogrammetric Engineering & Remote Sensing. 66: 1425–1435.

CHAPTER 5

Current Status and Recommendations for Future Work

Current status

This research shows that Douglas-fir, lodgepole pine, and ponderosa pine can be identified and mapped with reasonable efficacy in a mixed coniferous forest setting using a combination of lidar structure and intensity metrics. Western larch is not uniquely identifiable, consistently appearing in a 'mixed' class. Although the field data do not yet exist to conclusively document the performance of the ALS-based classifications at plot or stand-level for the mixed class, a comparison of predicted species distributions with walk-through inventory data provides compelling anecdotal evidence that the methodology proposed could provide inventory-quality species data along with the more conventional measurements of stand height, canopy cover, and their derivatives. Species prediction at tree-scale is more uncertain with low-density datasets like the ones used in this study, given uncertainties in the stem location data. More likely, individual trees will need to be attributed with species data from high-quality pixel-based or stand-based species distributions like the ones provided by the MLC classification, a possibility that provides fertile ground for additional research.

The original contributions of the work described herein are three-fold. First, the research moves the literature beyond the frequently used discriminant analysis at plot-scale to classify and validate species distributions across a relatively large landscape. Second, the Maximum Likelihood Classification using modified LDA output with percent canopy cover is novel, simple to apply, provides good results, and is easily modifiable to accommodate additional variables or to develop a priori probabilities based on independent species distribution models. Third, results are provided for a new suite of species on a landscape not previously investigated. Additionally, the effects Automatic Gain Control, scan angle, and range are evaluated, one of four attempts to do so to date in the literature.

Results of this research are encouraging and open up the possibility of improving the methodology used for tree species discrimination to support various management activities, including timber production, biomass estimation, wildlife habitat delineation, and site productivity prediction. The system used for this study, Leica ALS50, operates with automatic gain control (AGC) to record target intensity. This results in the automatic adjustment of captured raw intensity values for system gain, slant range, and flying height variations (Leica Geosystems, 2008). Unfortunately, there are only few studies focusing on such adjustments (Adams, 2000; Wagner et al, 2004; Korpela, 2008) and they do not corroborate whether it is consistent for every application for similar targets. Indeed, the AGC concept is considered a nonlinear system, and it is always difficult to solve for an effect arising from the adjustment of an appropriate nonlinear equation (Martinez, 2001). Therefore, this study has had to deal with two issues generated by the system: (1) lack of detailed insight into the proprietary pulse detection algorithms and (2) an uncertain AGC adjustment. While these problems may be sufficiently acknowledged, a lack of reflectance measurements of the targets in the field/laboratory using a spectrometer make it difficult to evaluate and compare the signal strength values produced by the system. Indeed, admittedly, it becomes more complicated to calculate the approximate saturation of minimum and maximum values of 8-bit resolution used in the ALS50 intensity conversion. For example, the removal of 0 and 255 values utilized in this study is based on previous observations of ALS datasets and there is no documentation to corroborate them.

For species discrimination at individual tree level, it is important to recognize that the stem identification algorithm used in this study produced a root mean square error (RMSE) of 17 stems per all plots (46.8%) for overstory and intermediate trees across all structure types (Rowell et al., 2009). With such RMSE, there is a potential error in identifying trees using the methodology described in Chapter 3 due to a an uncertain number of trees belonging to the plots used in this analysis. Additionally, the algorithm also generates higher RMSE (21 stems) for regeneration trees (height <6 meters), including understory trees (Rowell et al., 2009). The inclusion of these classes into taller trees of similar species may affect classification accuracy. Therefore, the methodologies described in this research should be applied with caution to individual trees. Indeed, there is not an attempt to perform error propagation for this part of the study due to the assumptions used in the methodology outlined in Chapter 3 that identified coincident trees in both laser and field data.

The Linear Discriminant Analysis (LDA) and Maximum Likelihood Classification (MLC) are based on the assumption that all classes are normally distributed. However, it is acknowledged that it was not possible to fulfill this requirement due to the small number of samples for dominant species, especially for western larch and lodgepole pine. As depicted in the box-whisker plots in Chapter 3, some classes are not distributed normally and several transformations have been performed to overcome this problem with insignificant changes to results. Additionally, Groeneveld (1991) and Zuuring (2009) suggest that absent considerable skewness, the data should be sufficient for statistical testing. Such assumption is also applied to the variables used in MLC for mapping species distribution at landscape level described in Chapter 4.

Despite these potential problems, the use of airborne laser scanning data for species identification is more promising than ever and the challenges encountered in this research provide important context for future work in the domain of species mapping with ALS data.

Future Work

Based on the current results, several recommendations for future studies are proposed:

1. Compare the results of presented methodology with the results of alternative methodologies. For example, a conventional LDA using all functions may produce more efficient computation and better results. Alternatively, the twelve input variables could be examined within other classification routines such as MLC and perhaps the variables could be reduced using principle component analysis. In point fact, many of these alternatives were explored before arriving at the current methodology, but they were not explored comprehensively or exhaustively.
2. Develop a methodology to attribute individual trees from pixel-based or stand-based species information. The methodology could be based on simple probabilities of occurrence from the pixel data or could include knowledge of growth habit of individual species. For example, understory trees are not likely to be shade-intolerant

ponderosa pine, western larch or lodgepole pine. More likely, they are Douglas-fir, Engelmann spruce, or sub-alpine fir. The *a priori* probabilities used in the Maximum Likelihood Classification provide flexibility to deal with uncertainty for species assignment.

3. The incorporation of species distributions information based on site-species relationships is expected to improve classification by predefining the places a certain species can occur. Again, this could be performed using *a priori* probabilities in the Maximum Likelihood Classification.
4. Instead of using a single standard deviation to define species classes within an LDA function, incorporating modified standard deviations for each class may improve classification accuracy. Examining the effects of different class centroids and thresholds may prove beneficial in further analysis.
5. The application of a higher density ALS data is likely to improve the classification accuracy at the individual tree scale by improving stem identification and by better characterizing the intensity and structural variability of individual trees.
6. There are many uncertainties regarding automatic gain control (AGC), and in many ways, raw intensities may be preferable to AGC-corrected intensity because of the ability to control normalization in more precisely. Intensity data are still not a highly quantitative and there is significant room for improvement. Additionally, the proprietary behavior of lidar vendors is a chronic problem that needs constant attention to effect change.
7. More field plots that better represent species distributions would improve the analysis, particularly for under-represented species like western larch.

8. A combination of aerial imagery with airborne laser scanning data may improve the results of this study, especially to separate non-mixed species classified into mixed classes. For example, the unique phenology of western larch could be exploited to separate it from other species.

References:

Adams, M.D., 2000. Lidar design, use and calibration concepts for correct environmental detection. IEEE Transactions on Robotics and Automation,16(6), 753-761.

Groeneveld, R.A., 1991. An influence function approach to describing the skewness of a distribution. The American Statistician, 45, 97-102

Martinez, I.G., 2001. Automatic Gain Control (AGC) circuits: Theory and design. Term paper. http://www.eecg.toronto.edu/~kphang/papers/2001/martin_AGC.pdf (accessed on March 16, 2009).

Korpela, I.S., 2008. Mapping of understory lichens with airborne discrete-return LiDAR data. Remote Sensing of Environment 112 (10), 3891–3897.

Leica Geosystems, 2008. ALS50 Airborne Laser Scanner guide. Norcross, Georgia. 10pp.

Rowell, E., Seielstad, C., Queen, L., 2009. Implementation and Performance of a Local-Maximum Tree Stem Identification Algorithm Using Vector-format Airborne Laser Scanner Data. Manuscript in preparation.

Wagner, W., Ullrich, A., Melzer, T., Briese, C., Kraus, K., 2004. From single-pulse to full-waveform airborne laser scanners: potential and practical challenges. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 35 (Part B3), 201–206.

Zuuring, H., 2009. Personal communication. The University of Montana.

Appendix A

DATA RECORDS FROM HORIZONS, INC.

The National Center for Landscape Fire Analysis of the University of Montana's College of Forestry and Conservation received two groups of DVDROMs containing (1) raw preprocessed lidar data in binary .las format, and (2) log files and documentation.

1. The "Preliminary Preprocessed LAS Data (Unbiased)" were stored on eight DVDs. Seven DVDs hold data acquisitions conducted on July 27- 29, and August 17, 2005 for Lubrecht Experimental Forest and Missoula International Airport, and One DVD contains an additional data acquired on June 29, 2006 for Elk Creek watershed area.

   File names:

   a. Missoula International Airport

   LDR060621_150128_1.LAS
   LDR060621_144451_1.LAS
   LDR060621_145625_1.LAS
   LDR060621_151943_1.LAS

   b. Lubrecht Experimental Forest

   LDR050616_144514_1.LAS
   LDR050616_154124_1.LAS
   LDR050619_192627_1.LAS
   LDR050616_155212_1.LAS
   LDR050619_192159_1.LAS
   LDR050619_174254_1.LAS
   LDR050619_191050_1.LAS
   LDR050619_182128_1.LAS
   LDR050619_185338_1.LAS
   LDR050616_150104_1.LAS
   LDR050616_152743_1.LAS
   LDR050619_175242_1.LAS
   LDR050619_183741_1.LAS
   LDR050619_190158_1.LAS
   LDR050616_143049_1.LAS
   LDR050619_181120_1.LAS
   LDR050616_141909_1.LAS
   LDR050619_183247_1.LAS
   LDR050619_184410_1.LAS
   LDR050619_173329_1.LAS
   LDR050616_142357_1.LAS
   LDR050616_150952_1.LAS
   LDR050616_153746_1.LAS
   LDR050616_155736_1.LAS
   LDR050619_180121_1.LAS
   LDR050616_143749_1.LAS
   LDR050616_151849_1.LAS
   LDR050616_145220_1.LAS
   LDR050619_182724_1.LAS

c. Elk Creek watershed:

LDR060621_145155_1.LAS
LDR060621_142740_1.LAS
LDR060621_133217_1.LAS
LDR060621_152840_1.LAS
LDR060621_124809_1.LAS
LDR060621_121124_1.LAS
LDR060621_134009_1.LAS
LDR060621_131804_1.LAS
LDR060621_135549_1.LAS
LDR060621_141435_1.LAS
LDR060621_124140_1.LAS
LDR060621_150746_1.LAS
LDR060621_123400_1.LAS
LDR060621_130935_1.LAS
LDR060621_122800_1.LAS
LDR060621_152448_1.LAS
LDR060621_130241_1.LAS
LDR060621_125528_1.LAS
LDR060621_140824_1.LAS
LDR060621_140017_1.LAS
LDR060621_141919_1.LAS
LDR060621_132528_1.LAS
LDR060621_135059_1.LAS

2. The contents of the documentation DVD include:

   a. Raw Aero files containing raw scanned images during acquisition
   b. Global Positioning Systems (GPS)
   c. Intertial Measuring Unit (IMU)
   d. Log Files of acquisition settings
   e. Flight Logs

Appendix B

ACQUISITION SETTINGS PROVIDED BY HORIZONS, INC.

Horizons, Inc. conducted LIDAR data acquisitions requested by the National Center for Landscape Fire Analysis (NCLFA) for the Lubrecht Experimental Forest and Elk Creek watershed, Montana. The project was carried out on July 27- 29 and August 17, 2005 and June 29, 2006 using Leica ALS50 operated with automatic gain control (AGC).

The acquisitions included the survey control at Missoula International Airport (MSO) for determining position and orientation references. The coordinate position of MSO was **46 55 05.23687 N** and **114 05 36.30852 W** at the ellipsoid height of **960.03 meters**.

Acquisition setting for MSO:

Flying Height	750 meters AMTE	1250 meters AMTE
Airspeed	72.02 meters/sec	72.02 meters/sec
Laser Pulse Rate	43 KHz	43 KHz
Field of View (FOV)	35°	35°
Scan Rate	35 Hz	35 Hz
Average Swath Width	473 meters	788 meters
Estimated Signal to Noise		
-for 10% diffuse targets	30.35 at nadir	31.94 at nadir
	28.74 at FOV edge	30.22 at FOV edge
-for 5% diffuse targets	20.66 at nadir	21.84 at nadir
	19.50 at FOV edge	20.61 at FOV edge

Acquisition setting for LEF and Elk Creek watershed:

Flying Height	1829 meters AMTE
Airspeed	72.02 meters/sec
Laser Pulse Rate	36 KHz
Field of View (FOV)	35°
Scan Rate	27 Hz
Average Swath Width	1153 meters
Estimated Signal to Noise	
-for 10% diffuse targets	12.34 at nadir
	11.54 at FOV edge
-for 5% diffuse targets	7.73 at nadir
	7.73 at FOV edge

Flight Plan at LEF and Elk Creek watershed

LINE ID	LAT	LON	ALT	% RTN 1	% RTN 2	% RTN 3	INTENSITY	AGC
060621_121124	46.960394	-113.558	2921.306	100	18	2.6	157	123
060621_121124	46.839941	-113.236	2926.67	99.8	48.6	11.8	162	120
060621_122800	46.868532	-113.249	3614.441	100	61	11.8	96	125
060621_122800	46.895867	-113.25	3612.519	100	46.2	6.8	101	126
060621_123400	46.895207	-113.258	3623.894	100	39.2	8.4	117	124
060621_123400	46.821777	-113.254	3628.18	100	25	4	149	123
060621_124140	46.820265	-113.262	3646.66	100	28	5.2	145	123
060621_124140	46.90405	-113.266	3638.346	100	39.4	6	119	124
060621_124809	46.903539	-113.273	3535.057	100	50.4	9.6	118	123
060621_124809	46.80194	-113.269	3549.32	100	25.8	4.6	150	123
060621_125528	46.791054	-113.276	3475.9	100	14.6	1.8	179	122
060621_125528	46.91391	-113.281	3485.546	100	46.2	9.2	124	123
060621_130241	46.913347	-113.289	3464.554	100	37.4	7	136	123
060621_130241	46.791938	-113.284	3469.028	100	36.6	6.4	141	124
060621_130935	46.791164	-113.291	3400.397	100	19.8	3.4	180	121
060621_130935	46.911966	-113.296	3392.289	100	43.8	7.8	128	122
060621_131804	46.932451	-113.304	3365.042	100	31.8	6.2	144	123
060621_131804	46.791815	-113.299	3361.17	100	34.8	4.8	138	124
060621_132528	46.790798	-113.306	3335.644	100	26.6	5	172	122
060621_132528	46.930772	-113.312	3324.435	100	27.2	5.2	156	123
060621_133217	46.931529	-113.32	3245.494	100	15	1.2	183	121
060621_133217	46.800179	-113.314	3218.481	100	12.6	1.2	179	120
060621_134009	46.799619	-113.322	3228.656	100	42.6	8.2	129	124
060621_134009	46.932078	-113.327	3241.134	100	10.6	1.2	185	122
060621_135059	46.845259	-113.331	3613.211	99.8	41.6	10.8	102	128
060621_135059	46.807484	-113.33	3620.653	100	22.6	5.4	152	123
060621_135549	46.809247	-113.337	3545.763	100	32.8	7.2	129	124
060621_135549	46.844235	-113.339	3546.835	100	39.2	10.2	115	125
060621_140017	46.845053	-113.346	3479.568	100	36.8	6	118	125
060621_140017	46.817057	-113.345	3481.912	100	20.8	4	162	123
060621_140824	46.832199	-113.361	3323.055	100	29.4	6.4	156	122
060621_140824	46.834694	-113.241	3326.058	100	45.2	14.2	132	122
060621_141435	46.887268	-113.25	3231.874	100	56.2	9.8	120	122
060621_141435	46.885736	-113.337	3233.914	100	29.4	3.8	140	124
060621_141919	46.85022	-113.339	3233.539	100	33	4.4	143	123
060621_141919	46.85134	-113.24	3236.809	100	37	7.8	149	121
060621_142740	46.841239	-113.271	3544.102	100	21.4	3.2	155	123
060621_142740	46.901095	-113.273	3559.003	100	44.6	8.6	118	124
060621_144451	46.907913	-114.077	2224.102	99.4	1.2	0	158	125
060621_144451	46.926187	-114.108	2228.033	100	1.4	0	158	124
060621_145155	46.927476	-114.111	2223.93	100	0.6	0	149	125
060621_145155	46.906832	-114.076	2231.917	99.4	0	0	157	125
060621_145625	46.90909	-114.072	2225.76	99.8	0	0	163	125
060621_145625	46.928444	-114.104	2226.414	100	2.2	0.4	153	127
060621_150128	46.916814	-114.104	2215.773	100	0	0	159	126
060621_150128	46.927637	-114.09	2222.443	99.6	0.2	0	136	127
060621_150746	46.906609	-114.076	1728.802	99.6	0	0	154	130

LINE ID	LAT	LON	ALT	% RTN 1	% RTN 2	% RTN 3	INTENSITY	AGC
060621_150746	46.92625	-114.108	1719.378	100	0	0	157	129
060621_151943	46.926735	-114.109	1722.934	99.6	0.6	0	157	129
060621_151943	46.906673	-114.075	1724.499	100	0	0	154	129
060621_152448	46.908794	-114.073	1732.942	98.8	0	0	159	131
060621_152448	46.927383	-114.105	1727.136	100	1.6	0	147	131
060621_152840	46.914919	-114.105	1717.228	100	0	0	154	130
060621_152840	46.927644	-114.089	1721.671	99.8	0.4	0	133	131

Post flight processing

Horizons, Inc. performed all post flight processing, including the integration of IMU data that recorded changes in positions and orientations with GPS data, which recorded the target location. The LIDAR, IMU and GPS data were correlated using the recorded GPS time to determine the coordinate of each return. The Leica ALS50 can receive up to four returns per each laser shot allowing multiple returns from vertical objects as the laser reflected toward the ground, such as vegetation The "Preliminary Preprocessed LAS Data (Unbiased)" delivered to NCLFA were in binary .las format containing information of LIDAR return of each target a position, including three dimensional coordinates (**xyz**), **intensity, number of returns per shot, type of returns, system gain value (AGC), scan angle and GPS time**.

Bare Earth and Canopy Separation

The NCLFA performed bare earth and canopy separation using Microstation Development Language (MDL) extension in TerraScan. The LEF lidar coverage was divided into 78 tiles (1 km^2) with 50 meter overlap on tile borders and each tile was exported as ASCII.

The headers of x, y, z, R1,R2, AGC, scan angle were added using Programmer's File Editor (text editor). Following, the tiles were reprojected into NAD83 UTM Zone 12N. The canopy height was calculated using the 'tinspot' command in Arc by subtracting the canopy elevation (z) with bare earth created using triangulation irregular network (TINs) (resolutions of 0.5m, 1m, and 2m). The final products are two groups of lidar datasets, bare earth and canopy height model in shapefile format.